DISSERTATION

SPATIAL ANALYSIS OF HUMAN LYME DISEASE RISK IN AN ENDEMIC COUNTY

Submitted by

Kiersten Jenae Kugeler

Department of Environmental and Radiological Health Sciences

In partial fulfillment of the requirements

For the Degree of Doctor of Philosophy

Colorado State University

Fort Collins, Colorado

Summer 2014

Doctoral Committee:

Advisor: Jennifer Peel
Co-advisor: John Reif

Lars Eisen
Paul Mead
Brad Biggerstaff

ABSTRACT

SPATIAL ANALYSIS OF HUMAN LYME DISEASE RISK IN AN ENDEMIC COUNTY

An understanding of the factors that drive spatial variation in human Lyme disease risk is important for appropriate development and implementation of public health interventions. Yet, these factors are poorly understood. This dissertation utilized fine-scale environmental and human Lyme disease data from a single county to quantify the spatial distribution of human Lyme disease occurring 2001-2011 and to evaluate whether spatial variation in disease risk was explained by several factors, including land use, land cover, deer density, and tick infestation on deer.

All studies were conducted with data from Howard County, Maryland. The first project described spatial clustering of human Lyme disease according to residence. When compared to other areas of the County, areas with elevated disease risk were characterized by more low-density development and more red and white oak forest. The second project used multilevel (i.e., mixed effect) models to examine risk factors for human Lyme disease among all homes in Howard County. In this analysis, 8% of all variation in human disease risk was due to the census block group location of households; the remaining variation in human disease risk occurred within census block groups. Most of the variation in risk between census block groups was explained by household-level land use and land cover characteristics and census block group-level differences in forest and socio-demographics, yet some variation in risk between block groups remained unexplained with available covariates. Increased risk of Lyme disease was

associated with low- and medium-density residential development, red and white oak forest, increasing proportion of the census block group classified as forest, and residing in a census block group characterized by higher income, home value, and education. The third project evaluated associations between deer density, tick infestation on deer, and human disease risk. Study findings suggested that areas with lower deer density had higher abundance of ticks on deer and higher risk of human Lyme disease. These results suggest that moderate deer reduction in inland areas, as occurs through community deer management programs, may not be a viable Lyme disease prevention measure.

This dissertation advances knowledge of the fine-scale epidemiology of human Lyme disease and demonstrates the importance of using human outcome data, in addition to entomologic data, to understand variation in Lyme disease risk. These studies use advanced analytic methods to demonstrate significant sub-county spatial variation in risk of human Lyme disease, validate previously recognized risk factors for human illness, identify novel associations of a specific forest type with human disease, and demonstrate the importance of human behavior in placing humans at risk. Finally, results of this dissertation suggest that additional analyses using multilevel modeling techniques may help to provide insight regarding many remaining questions in the epidemiology of Lyme disease.

ACKNOWLEDGMENTS

I would especially like to acknowledge my advisor, Jennifer Peel. Without your knowledge, patience, and encouragement, this work would not have been possible. Under your mentorship, the knowledge I have gained through this degree-seeking process has been broader than I possibly imagined. To John Reif, my co-advisor—I started down this path in your office in the summer of 2009 and have appreciated your continual guidance since then.

To the remaining members of my committee: Lars Eisen—I appreciate your immense knowledge of the subject matter and your ecologic perspective on all things Lyme disease. Brad Biggerstaff—Thank you for your patience with my endless statistical questions and for the mentorship you provided along the way. To the last member of my committee, my supervisor, Paul Mead—I am grateful you gave me the chance to be an epidemiologist in the first place. I appreciate your leadership and friendship.

I would like to thank the colleagues with whom I have worked over the years—I have learned so much and enjoyed my job immensely because of all of you. Thank you all for your support—I would list you all individually, but you will not fit into the one page maximum I have for acknowledgments! To Alison Hinckley, you are a great colleague and friend—I would not have set off down this path if not for you. Thank you for the pep talks! Moreover, I thank Andrias Hojgaard and Martin Williams for performing all of the tick pathogen testing. Finally, I am indebted to Phil Norman and Katherine Feldman in Maryland.

DEDICATION

I dedicate this dissertation to family; without your unconditional support, this dissertation would not have been possible. To my parents—you have always had confidence in me. To my sisters—our sisterly love and antics constantly make me smile. To my daughters, Britt and Teagan—someday you will understand why mom went to “work” and “school” and missed an occasional good night kiss. I hope one day this journey of mine gives you assurance that you can do anything as long as you work hard. To my husband, Brian—you are my biggest supporter and my best friend. Without your love, support, endless encouragement, and fantastic parenting, I would never have started down this road, let alone reached the end.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1

1. INTRODUCTION AND BACKGROUND

INTRODUCTION

Summary of literature and rationale for dissertation

Lyme disease is a multisystem zoonotic illness caused by spirochetes of the *Borrelia burgdorferi* sensu lato complex and transmitted to humans by *Ixodes* species ticks (Figures 1.1. and 1.2.). The disease occurs in specific geographic areas of North America, Europe, and Asia that support the enzootic cycle of the bacteria. In the United States, Lyme disease is the most common vector-borne disease and typically the sixth most common nationally notifiable disease (Centers for Disease Control and Prevention 2013a). Lyme disease does not occur uniformly across the United States; 13 states in the Northeast, mid-Atlantic and north-central regions account for approximately 95% of all reported cases (Figure 1.3.)(Steere et al. 1984; Nadelman and Wormser 1998; Steere 2001; Centers for Disease Control and Prevention 2013a).

The incidence and geographic range of Lyme disease have increased consistently since its discovery in the late 1970s in Lyme, Connecticut, despite development of several possible prevention methods (Figure 1.4.)(Stafford 2007; Centers for Disease Control and Prevention 2013a). In the absence of a commercially-available vaccine, there is clear need for better prevention options if the disease is to be controlled in the United States. Theoretically, prevention of Lyme disease can occur at the individual level (e.g., tick checks, repellent use), household level (e.g., landscape modification to reduce likelihood of infected ticks in the

peridomestic habitat), or community level (e.g., deer population reduction). Community-level interventions may hold the most potential because of their broader spatial impact and because they do not rely on individual human behavior to achieve success. Robust information regarding risk factors for human disease is critical not only for the continued development of novel prevention methods, but also for appropriate, effective, and successful implementation of prevention methods in high-risk areas.

The risk of acquiring Lyme disease depends upon the abundance of infected ticks and on human behavior patterns that facilitate interaction with tick-infested habitat. Few environmental risk factors (defined for the purposes of these projects as factors other than human behavioral factors that increase risk of human disease) have been identified consistently; proximity to forest and low-density residential development are most commonly associated with disease (Killilea et al. 2008). The underlying biologic mechanism for these associations is not well understood but may be related to elevated potential for human contact with forested and forest edge environments in which ticks, rodent reservoir hosts, and white-tailed deer all coexist (Jackson et al. 2006a; Killilea et al. 2008). Nevertheless, tick abundance and infection prevalence vary across space not only between regions, but also between individual properties within endemic areas (Maupin et al. 1991; Ostfeld et al. 1995; Nicholson and Mather 1996; Kitron and Kazmierczak 1997; Wilson 1998; Pardanani and Mather 2004; Connally et al. 2006). Environmental risk factors should be refined beyond simply “forest” in order to better target environmental or behavioral interventions in a meaningful way. A more specific understanding of environmental risk factors could help develop and guide appropriate implementation of public health interventions.

The spatial scale of analysis is important when examining the association of explanatory factors with Lyme disease risk. Most analyses of environmental risk factors for human Lyme disease have been conducted at coarse spatial scales—at the county level and above. This approach is helpful when focusing prevention and education resources in specific states and counties but masks sub-county variation and limits our overall understanding of disease risk. The spatial variation in human Lyme disease risk has not been well quantified on a fine (i.e., sub-county) spatial scale. Additionally, there is limited information regarding risk factors that act on the scale of an individual residential parcel or neighborhood.

This body of work examines the spatial variation in human Lyme disease risk and environmental and socio-demographic factors that may be associated with spatial variation in risk, all within an endemic county. The findings from these studies may help refine understanding of factors associated with human Lyme disease and consequently help inform future directions for Lyme disease prevention research.

Dissertation overview

The studies described in this dissertation examine the spatial distribution of human Lyme disease in one county (Howard County, Maryland) and evaluate whether spatial variation in disease risk is associated with several factors, including land use, land cover, deer density, and tick abundance on deer. The remainder of Chapter 1 provides general background information regarding Lyme disease, specifically its history, ecology, transmission, clinical manifestations and human disease surveillance. Additionally, Chapter 1 contains a more detailed review of literature relevant to the spatial distribution of and environmental risk

factors for Lyme disease, including deer density and tick abundance on deer, in the eastern United States. Chapter 2 (Project 1) describes the first known effort to quantify the clustering of human Lyme disease on the sub-county level, using the residential address of patients with Lyme disease reported to the public health system as case houses and all other households without reported Lyme disease as non-case households. Chapter 3 (Project 2) describes the assessment of environmental and socio-demographic factors associated with Lyme disease using an historical cohort of all households in Howard County. The multilevel analysis includes household location-specific environmental explanatory factors, as well as area-based forest fragmentation and census block group socio-demographic variables. Chapter 4 (Project 3) examines the association of sub-county variation in deer density with human Lyme disease risk. This last analysis is augmented by inclusion of measures of tick abundance on deer, which may mediate the effect of deer density on human disease risk. The study population in Project 3 is a subset of households in Howard County, those within a half-mile of parks with robust deer density measurements and hunting data. All studies use fine-scale exposure and outcome data from one county with an endemic pattern of Lyme disease incidence. These studies were conducted at a spatial scale relevant to feasible implementation of possible prevention methods. The remaining chapters contain a summary of collective findings and recommendations for future research, references, and appendices containing supplemental analyses for each project.

The State of Maryland consistently reports a high incidence of human Lyme disease. All studies described herein use data from Howard County, Maryland (located 10 miles from Baltimore and 20 miles from Washington D.C.; Figures 1.5. and 1.6.)(Bacon et al. 2008; Centers

for Disease Control and Prevention 2013a). The approximately 252 square miles had a population in 2010 approximately 287,000 with a population density of 1,145 residents per square mile (US Census Bureau 2012). The eastern half of the county has light urban and suburban development; the western part of the county transitions to primarily rural and agricultural uses. This county was selected as the study location due to the emergence of Lyme disease during the study time frame (2001-2011) and the availability of detailed geographic and environmental information, particularly deer density and tick abundance on deer information.

In this series of projects, fine-scale environmental data and geocoded address data in Howard County is used to: 1) describe and quantify spatial and temporal clustering of human Lyme disease cases and describe environmental and socio-demographic factors associated with high-risk clusters (Project 1); 2) evaluate the association between several environmental and socio-demographic factors assessed at the household and neighborhood level with human Lyme disease risk (Project 2); and 3) evaluate the associations between other environmental variables, specifically deer density and tick abundance on deer, and spatial variation in human Lyme disease risk among a subset of Howard County households (Project 3).

BACKGROUND

History of Lyme disease

Lyme disease was first identified during an investigation of a cluster of juvenile arthritis patients in Old Lyme, Connecticut, in the late 1970s (Steere et al. 1977). A few years later, the etiologic agent, *Borrelia burgdorferi,* was discovered (Burgdorfer et al. 1982; Steere et al. 1983). Lyme disease occurs in parts of the northern hemisphere; however, the etiologic genospecies, vector species, ecology and clinical manifestations vary slightly across continents.

Although the spirochete and its associated clinical syndrome were identified a few decades ago, there is evidence of its existence in Europe and North America long before. *B. burgdorferi* was identified in museum specimens of ticks collected in the late 1800s in Europe and early 1900s in the northeastern United States (Persing et al. 1990; Matuschka et al. 1996). A clinical syndrome reminiscent of Lyme disease was first described by a Swedish physician in 1909 (Afzelius 1910); a similar rash associated with a tick bite was described in Wisconsin in 1969 (Scrimenti 1970).

Since the discovery of Lyme disease as a distinct clinical entity, the incidence and geographic distribution of the disease have increased dramatically (Figures 1.3. and 1.4.). This increase may be due to enhanced detection and reporting as well as geographic expansion and a true increase in disease risk (Bacon et al. 2008). Currently, Lyme disease is the most common vector-borne disease in the United States, with approximately 30,000 confirmed and probable cases reported to the public health system annually (Centers for Disease Control and Prevention 2013a). In 2011, Lyme disease was the sixth most common nationally notifiable disease; during the same year, it was the second or third most common illness reported in highly endemic states (Centers for Disease Control and Prevention 2013a).

Ecology of Lyme disease

Lyme disease spirochetes are transmitted to humans by *Ixodes* species ticks. *Ixodes scapularis* (the blacklegged tick or "deer tick" [also referred to as *Ixodes dammini*]) is the vector in the Northeast and north-central United States (Figures 1.1. and 1.2.)(Steere et al. 1978; Steere and Malawista 1979). Lyme disease also occurs sporadically along the Pacific Coast of

the United States, where the human vector is *Ixodes pacificus.* In Europe and Asia, the vector species are *Ixodes ricinus* and *Ixodes persulcatus*, respectively.

Ticks are not insects; they are arthopods closely related to mites and spiders (Class: Arachnida, Subclass: Acari)(Goddard 2003). Ticks require blood for development, and are efficient vectors of many bacterial, viral, and protozoan agents. When actively searching for a host, ticks that vector *B. burgdorferi* "quest" on vegetation. They ascend vegetation, such as a blade of grass, and wait for an animal to approach. Signals such as vibration or carbon dioxide may influence ticks' readiness to climb on a passer-by (Goddard 2003). Based on morphological and behavioral characteristics, ticks are classified as hard ticks (Ixodidae), soft ticks (Argasidae), or Nutellidae*,* a rare group that have features of both hard and soft ticks (Goddard 2003). The *Ixodes* species ticks that vector *Borrelia burgdorferi* are hard ticks with three life stages. In a roughly two-year life cycle, *I. scapularis* consume three blood meals, one each as larvae, nymphs, and adults (Figures 1.1. and 1.2.)(Piesman and Spielman 1979; Yuval and Spielman 1990). Following a successful blood meal, adult females drop off the animal host and lay eggs, completing the life cycle.

The enzootic cycle of *B. burgdorferi* is a complex interaction between mammalian hosts and tick vectors (Figure 1.1.). Lyme disease foci are comprised of numerous host, vector, and environmental features, only some of which have been fully described. Transovarial transmission of *B. burgdorferi* occurs minimally (Piesman et al. 1986). Consequently, larval ticks only rarely transmit *B. burgdorferi* to humans; however, they may acquire infection during the first blood meal. Once acquired, infection persists transstadially (across life stages), presumably

for the remainder of the tick's life (Piesman and Sinsky 1988; Mather and Mather 1990; Mather et al. 1990). The white-footed mouse (*Peromyscus leucopus*) is generally considered the principal reservoir host for the enzootic cycle of *B. burgdorferi* in the eastern United States; the mouse is capable of maintaining high levels of spirochetemia, is a favored host for *I. scapularis* ticks, and can efficiently infect feeding ticks (Levine et al. 1985; Donahue et al. 1987). Other small rodent species, including shrews and chipmunks, as well as some bird species, also serve as reservoir hosts; these alternate species may be more important to the enzootic cycle than mice in some areas (Anderson and Magnarelli 1984; Mather et al. 1989; Telford III et al. 1990; Slajchert et al. 1997; Piesman and Gern 2004; Brisson et al. 2008).

White-tailed deer (*Odocoileus virginianus*) are the preferred blood meal host for adult *Ixodes* ticks but they are not infectious to feeding ticks (Piesman et al. 1979; Anderson and Magnarelli 1980; Main et al. 1981; Telford et al. 1988; Wilson et al. 1990b; Kurtenbach et al. 1998). Deer are important to the Lyme disease enzootic cycle because they host large populations of ticks in the reproductive stage. They are also critical to broader geographic dispersal of ticks because they travel farther than mice and small rodents. Adult blacklegged ticks feed on large mammals other than deer, although an estimated 95% of all adult feeding takes place on deer (Wilson et al. 1990b). Additional detail on the role of deer in the enzootic cycle is presented in the background section, "Deer density and ticks on deer as environmental risk factors for Lyme disease" (page 26).

Nymphal stage ticks are primarily responsible for the transmission of *B. burgdorferi* to humans. As a result of their small size, nymphs frequently go unnoticed for the several days of

tick attachment required to transmit *B. burgdorferi* (Figure 1.2.). Additionally, nymphal *I. scapularis* actively seek blood meals when humans are most active outdoors: late spring and early summer (Figure 1.1.)(Piesman et al. 1987; Falco et al. 1999). Moreover, this seasonal timing coincides with the occurrence of the vast majority of human cases (Bacon et al. 2008; Centers for Disease Control and Prevention 2013a). Adult ticks also transmit *B. burgdorferi* to humans, although fewer human cases occur during seasons when adult blacklegged ticks are active (fall and early spring)(Piesman et al. 1991; Falco et al. 1996; Bacon et al. 2008).

Although *I. scapularis* is found throughout much of the eastern United States, Lyme disease is not endemic in the southern or central states. Immature *I. scapularis* in the South prefer to feed on lizards rather than small rodents. Importantly, some lizards are inherently resistant to infection with *B. burgdorferi* and therefore do not infect immature ticks. This difference in host preference lowers the overall infection rate in the region (Apperson et al. 1993; Oliver et al. 1993; Piesman and Gern 2004).

Both biotic and abiotic factors contribute to the ecology of *I. scapularis* ticks in eastern North America (Kahl et al. 2002; Piesman 2002). These blacklegged ticks are unquestionably forest inhabitants. Tick distribution has been positively associated with deciduous, dry-mesic forests and sandy (or loamy sand) soils overlying sedimentary rock and negatively associated with grasslands, wet or pine forests, acidic and clay soils (Guerra et al. 2002). Blacklegged ticks thrive on humidity and desiccate easily, but need moist, rather than water-logged, settings to survive (Dister et al. 1997; Guerra et al. 2002; Piesman and Gern 2004). Deciduous forests provide ample leaf litter to protect ticks from desiccation and snowfall (Schulze and Jordan

1995; Piesman and Gern 2004). Although populations of *I. scapularis* have been associated primarily with oak and maple forests in the Northeast United States, they can inhabit a broad range of ecologic settings, particularly shrub-dominated habitat (Piesman and Spielman 1979; Maupin et al. 1991). Density of infected blacklegged nymphs has been associated with lower elevation, lower vapor pressure deficit, and low seasonal extremes in minimum temperature in a predictive acarological risk model in the eastern United States (Diuk-Wasser et al. 2012).

In the Northeast United States, Lyme disease is widely assumed by the scientific and medical community to be acquired primarily in the peridomestic (residential) environment (Pardanani and Mather 2004; Connally et al. 2006; Killilea et al. 2008; Eisen et al. 2012). This assumption is based on early studies that demonstrate presence of infected ticks on individual residential properties in endemic areas and recall of infected individuals (Falco and Fish 1988; Maupin et al. 1991; Nicholson and Mather 1996; Cromley et al. 1998; Orloski et al. 1998). However, the proportion of human Lyme disease cases that are acquired in the home environment is unknown and difficult to ascertain (Eisen and Eisen 2007; Killilea et al. 2008). Human movement, a long incubation period (i.e., 3-30 days, longer for those without a recognized erythema migrans rash) and frequent lack of recognition of tick bites complicate precise identification of an exposure location for most cases.

Pathogen transmission, clinical manifestations, and diagnosis

Borrelia burgdorferi is a motile spirochete with a small linear chromosome (approximately 1.5 MB) and several linear and circular plasmids (Fraser et al. 1997; Casjens et al. 2000; Steere et al. 2004). Virulence factors are not well-defined; however, several antigenic

proteins exist and form the basis for current serologic tests. The spirochete is proficient at host adaptation; several outer surface proteins (Osps) are expressed differentially within sites inside the tick and the mammalian host, coincident with variation in pH and temperature (Bergstrom et al. 2002). *Borrelia burgdorferi* sensu lato includes three genospecies that cause human Lyme disease (*Borrelia burgdorferi* sensu stricto, *Borrelia afzelii* and *Borrelia garinii*) and more than a dozen other organisms whose relevance as human pathogens is unknown (Barbour 1998; Stanek and Reiter 2011).

The Lyme disease spirochete disseminates locally in the skin at the site of the tick bite, initiating a robust innate immune response in which macrophages engulf and kill spirochetes (Steere et al. 2004). An erythema migrans (EM) rash at the site of the bite is an early sign of illness and is present in 70%-80% of patients (Smith et al. 2002; Steere and Sikand 2003; Bacon et al. 2008). The rash usually appears between three and 32 days following a tick bite, allowing differentiation between EM and allergic reactions to tick saliva; the latter occur within hours but typically subside within a few days. An EM rash is also usually accompanied by "flu-like" symptoms. EM can occur as a single lesion, or less commonly, as multiple lesions following dissemination in the bloodstream or lymphatic system (Sikand and Muellegger 2011). Without antibiotic treatment, the pathogen disseminates and can affect multiple organ systems including the heart, joints and nervous system (Steere 2001). Within days to weeks after illness onset, the organism has been documented in myocardium, retina, bone, spleen, retina, meninges and brain (Duray and Steere 1988). The spirochete invades and migrates through tissues, adheres to several types of host cells, and can evade immune clearance by changing antigenic expression (Steere et al. 2004).

Few organisms are present in infected individuals, limiting direct detection of infection; culture of EM rash biopsies and blood have low sensitivity (Aguero-Rosenfeld et al. 2005). Early Lyme disease is diagnosed clinically in areas where the disease is common; all other clinical circumstances require accompanying serologic tests (Wormser et al. 2006). A two-tier serologic approach is recommended. In this testing algorithm, a highly sensitive first tier enzyme immunoassay (EIA) is used as a screening test; positive or equivocal samples are then subjected to more specific western immunoblots to assess antibody reactivity to individual proteins (Centers for Disease Control and Prevention 1995; Wormser et al. 2006). Although recommended two-tier serologic tests are highly sensitive and specific for illness of duration longer than 4-6 weeks, sensitivity is low in early illness (Bacon et al. 2003; Aguero-Rosenfeld et al. 2005). Serologic assays can remain positive for months to years following successful treatment (Feder et al. 1992; Aguero-Rosenfeld et al. 2005). Approximately 7% of infections are asymptomatic (Steere et al. 1998; Steere et al. 2003).

Early antibiotic treatment can prevent more serious manifestations of illness (Wormser et al. 2006). Treatment of Lyme disease occurs with specific oral antibiotics in early disease, but may require intravenous therapy if not initiated until later in the course of illness (Wormser et al. 2006). Recognized death due to Lyme disease is rare (Kugeler et al. 2011). However, several cases of sudden death with post-mortem evidence of *B. burgdorferi* in cardiac tissue were recognized recently (Centers for Disease Control and Prevention 2013b). These findings suggest unrecognized death due to Lyme disease may be occurring. A post-infectious syndrome with clinical resemblance to chronic fatigue syndrome has been documented in a small percentage of patients (Nadelman and Wormser 1998; Wormser et al. 2006). The cause of these symptoms

is not understood, but may be a result of an abnormal immune response or persistent antigen stimulation from non-viable organisms (Steere et al. 2004; Chandra et al. 2010; Chandra et al. 2011a; Chandra et al. 2011b; Barbour 2012).

Human Lyme disease surveillance in the United States

Surveillance for human Lyme disease in the United States occurs according to state laws (Centers for Disease Control and Prevention 2013a). Public health officials of the Council of State and Territorial Epidemiologists (CSTE), the unifying organization of health departments, determine which conditions will be "nationally notifiable" and the standardized case definitions according to which potential cases will be classified (Teutsch and Churchill 2000; Centers for Disease Control and Prevention 2013a). Cases of notifiable conditions are reported voluntarily to the Centers for Disease Control and Prevention through the National Notifiable Diseases Surveillance System (NNDSS) (Teutsch and Churchill 2000; Centers for Disease Control and Prevention 2012; Centers for Disease Control and Prevention 2013a).

For these studies, human Lyme disease cases reported to the Howard County Health Department during 2001-2011 were used. The case definition for Lyme disease contains requirements for clinical presentation, laboratory findings, and exposure history. The 1996 CSTE case definition for Lyme disease was in effect for cases occurring during 2001-2007. Beginning in January 2008, a revised Lyme disease case definition was in place. As part of the 1996 case definition, in the absence of an EM rash, a confirmed case required laboratory evidence in the form of culture or demonstration of an antibody response. Two-tier testing was recommended but not required for confirmation (Council of State and Territorial Epidemiologists 1996). In the

2008 definition, the clinical and exposure criteria were unchanged. Criteria for laboratory evidence of infection were strengthened to require one of the following: 1) culture; 2) positive two-tier IgM serology, within the first 30 days of illness; 3) positive two-tier IgG serology; or 4) positive single-tier IgG western immunoblot (Council of State and Territorial Epidemiologists 2008). The other major revision in the 2008 case definition was allowance for probable and suspect cases, not solely confirmed cases. The effect of the case definition revision on reported case numbers has not been thoroughly examined; however, according to analyses performed by several state health departments, the total number of confirmed and probable cases classified under the 2008 definition is likely equivalent to the number of confirmed cases under the 1996 case definition (CDC, unpublished data).

The 2008 Lyme disease surveillance case definition (Council of State and Territorial Epidemiologists 2008) is excerpted here:

> …For purposes of surveillance, [erythema migrans (EM)] is defined as a skin lesion that typically begins as a red macule or papule and expands over a period of days to weeks to form a large round lesion, often with partial central clearing. A single primary lesion must reach greater than or equal to 5 cm in size across its largest diameter. Secondary lesions also may occur. Annular erythematous lesions occurring within several hours of a tick bite represent hypersensitivity reactions and do not qualify as EM. For most patients, the expanding EM lesion is accompanied by other acute symptoms, particularly fatigue, fever, headache, mildly stiff neck, arthralgia, or myalgia. These symptoms are typically intermittent. The diagnosis of EM must be made by a physician. Laboratory confirmation is recommended for persons with no known exposure. For purposes of surveillance, late manifestations include any of the following: musculoskeletal system--recurrent, brief attacks (weeks or months) of objective joint swelling in one or a few joints, sometimes followed by chronic arthritis in one or a few joints; nervous system--any of the following, alone or in combination: lymphocytic meningitis; cranial neuritis; radiculoneuropathy; or, rarely, encephalomyelitis; cardiovascular system--acute onset of high-grade (2nd-degree or 3rd-degree) atrioventricular

conduction defects that resolve in days to weeks and are sometimes associated with myocarditis.

Confirmed: a) a case of EM with a known exposure, or b) a case of EM with laboratory evidence of infection and without a known exposure or c) a case with at least one late manifestation that has laboratory evidence of infection.

Probable: any other case of physician-diagnosed Lyme disease that has laboratory evidence of infection

Suspected: a) a case of EM where there is no known exposure… and no laboratory evidence of infection, or b) a case with laboratory evidence of infection but no clinical information available (e.g. a laboratory report).

Under-reporting is inevitable in public health surveillance (Teutsch and Churchill 2000; Doyle et al. 2002). Under-reporting for Lyme disease has been estimated to be between three and 12-fold, meaning that for every one case reported, three to 12 Lyme disease cases actually occur and are diagnosed (Meek et al. 1996; Orloski et al. 1998; Naleway et al. 2002). Although the mechanisms of laboratory reporting and follow-up have changed over time and vary across jurisdictions, one component of the Lyme disease case definition has remained unchanged—direct reporting of EM rashes from clinicians in endemic areas. Recent research using clinical laboratory data including test volume, positivity rates for various tests, and published sensitivity and specificity of tests has yielded an estimate that roughly 300,000 human infections with *B. burgdorferi* occur in the United States each year (Hinckley et al. 2014). This estimate is ten-fold higher than the number of Lyme disease cases reported to the Centers for Disease Control and Prevention in 2011 (Centers for Disease Control and Prevention 2013a). Over-reporting is likely to occur as well, but may be less of a problem in areas where the disease is common than in areas where the disease is rare and the positive predictive value for Lyme disease laboratory tests is substantially reduced (Tugwell et al. 1997).

Current methods for Lyme disease prevention

Lyme disease has continued to increase in incidence and geographic range since its discovery in the late 1970s, despite research identifying several possible prevention and control options (Hayes and Piesman 2003; Bacon et al. 2008; Eisen et al. 2012; Centers for Disease Control and Prevention 2013a). In theory, prevention of Lyme disease is accomplished by interrupting the enzootic cycle with environmentally-based interventions or by implementing human-based behavioral interventions. Therefore, prevention can occur at multiple possible "levels": individual level (e.g., personal protective measures such as repellent use and daily tick checks), household level (e.g., landscape modification) or community level. Currently, two possible community-level interventions exist: deer population reduction and acaricide (tick pesticide) treatment of deer (using four-poster devices) (Hayes and Piesman 2003; Fish and Childs 2009).

Most efforts at preventing disease have focused on disseminating information and encouraging personal protective measures despite limited evidence regarding their effectiveness (Hayes and Piesman 2003). During the 1990s, human vaccines for Lyme disease were developed. A recombinant OspA vaccine with adjuvant was commercially available starting in 1999. In a large double-blind placebo controlled trial, efficacy for preventing Lyme disease was 49% after two injections and 76% after three injections (Steere et al. 1998). Due to rapidly waning antibody titers, booster doses were needed every one to three years. The Advisory Committee on Immunization Practices recommended that vaccination be considered in persons aged 15-70 years of age who live in or frequent areas of high risk for Lyme disease (Advisory Committee on Immunization Practices 1999; Steere 2001). Acceptance by both

practitioners and the public was limited, and the vaccine was withdrawn from the market in 2002. Some of the reasons for the limited vaccine demand were: low risk in most of the country, high cost and need for frequent booster injections, inability to provide vaccination to children (the highest risk group), and a misconception spread through advocacy groups that the vaccine triggered autoimmune arthritis (Steere 2001).

Other personal prevention measures include use of insect repellent, avoidance of tick habitat, wardrobe modification to improve tick detection and complicate tick attachment (wearing light-colored clothing and tucking pants into socks), checking the body daily for ticks, and showering within two hours of being outdoors. There are few studies in which the effectiveness of these measures has been assessed; evidence that these measures prevent human Lyme disease is limited (Ley et al. 1995; Orloski et al. 1998; Herrington 2004; Connally et al. 2009). The lack of evidence for the effectiveness of these measures may be both a result of difficulty in accurate quantification of their use and overall infrequency in use, even in areas where Lyme disease is highly endemic (Gould et al. 2008). Antibiotic prophylaxis following tick bite is another prevention option in endemic areas; however, many people with Lyme disease never recall a tick bite and thus do not seek care (Wormser et al. 2006).

Measures to prevent disease that can be implemented at the household level include landscape modification to minimize tick migration into the yard and treatment of the landscape with acaricide (Stafford 2007). Although a single springtime application of acaricide can kill most ticks for the duration of the summer in a non-residential setting (Schulze et al. 1991; Stafford 1991; Schulze et al. 2001b; Schulze et al. 2005; Stafford 2007; Schulze et al. 2008), a

recent well-powered randomized, placebo-controlled trial assessing this intervention found that even though acaricide killed ticks where it was applied, there was no reduction in tick encounters or human illness in the acaricide-treated households compared to placebo-treated households (CDC, unpublished data). Possible reasons for this discrepancy include: 1) acaricide cannot legally be applied near water, and thus killing ticks in certain areas does not necessarily translate to lower risk in the yard as a whole; 2) people are not acquiring ticks in their own yard; or 3) there is a threshold effect in the association between tick abundance and human-tick encounter risk. Nevertheless, this study suggests that area-wide application of acaricide by homeowners to prevent human illness reduced tick abundance but did not apparently reduce human risk. Further studies evaluating other intervention methods are underway.

Community-level measures to prevent Lyme disease are not in widespread use, but are potentially the most promising because of the broader spatial scale on which they intervene. Additionally, community-level measures do not rely on the practices of each individual to achieve success. Current theoretical options are acaricide application on deer and deer population reduction (Hayes and Piesman 2003). Interventions related to white-tailed deer have been suggested because deer are an important source of blood for adult blacklegged ticks. Acaricide treatment of deer (via four-poster feeding stations) may reduce tick populations by roughly 70%, after a delay of several years (Brei et al. 2009). The impact of four-poster devices on human disease has been assessed in only one study, with mixed results (Garnett et al. 2011). Widespread use of four-posters is hampered by cost and need for consistent personnel and resources to achieve success. Deer elimination or near elimination in isolated (i.e., island) settings may be a viable measure to interrupt the enzootic cycle of Lyme disease. Evidence for

more moderate deer reduction (not nearing elimination) as a viable prevention method is limited (Stafford 2007). A few intervention studies have examined the impact of deer reduction or elimination on blacklegged tick abundance, mostly in island or isolated geographic settings not representative of the mainland ecology where most persons at risk for Lyme disease reside (Wilson et al. 1984; Wilson et al. 1988; Deblinger et al. 1993; Stafford et al. 2003; Rand et al. 2004; Jordan et al. 2007). These studies have shown that when deer are eliminated or nearly eliminated, tick reproduction and abundance are substantially reduced, but the relationship may not be linear. The effect of deer reduction in an inland setting open to deer immigration is unclear.

Spatial variation in risk of human Lyme disease

Description of spatial patterns in disease occurrence has been intrinsic to the practice of epidemiology since its infancy (Snow 1855; Palm 1890). Description of the distribution of health outcomes necessarily includes not only "who" is affected but "where" a condition of interest occurs. The geographic distribution of a health outcome itself may provide clues as to the source of the disease (Waller and Gotway 2004). If a disease displays geographic clustering, the cluster size itself can reveal a scale at which control measures and education efforts are most appropriate (Glavanakov et al. 2001). A growing body of literature devoted to spatial statistics and spatial patterns of disease is linked to terms such as "spatial epidemiology" or "health geography". Use of geographic information systems (GIS) and spatial statistical models to study environmental risk factors for vector-borne diseases has increased in recent years (Kitron 1998; Eisen and Eisen 2011).

Spatial variation in entomologic measures of tick abundance and infection prevalence occurs not only across states and regions but between locations only a few hundred meters apart (Pardanani and Mather 2004; Killilea et al. 2008). In contrast, analyses of spatial and temporal variation of human Lyme disease cases have been reported at coarse spatial scales (at the county level and above). Specifically, Waller et al. (2007) used hierarchical modeling methods to describe spatial and temporal patterns of reported Lyme disease across states and counties in the Northeast United States during 1990-2000 and described regional trends. Spatial and temporal heterogeneity of county-level incidence in New York State has been described by multiple authors (Glavanakov et al. 2001; Chen et al. 2005; Chen et al. 2006). Glavanakov et al. (2001) found that although the Lyme disease epidemic advanced geographically during 1986-1996, there was a consistent positive spatial autocorrelation in incidence rates within a 120 km distance of the southeastern New York focus. In other words, counties with centroids within this distance of each other experienced similar trends in incidence over time. Likewise, Chen and colleagues measured spatial autocorrelation of county incidence rates in New York during 1990-2000 and found similar results. They also described changing trends in demographic characteristics of reported cases during the study period (Chen et al. 2005). The authors noted that although incidence rates increased overall during the study period, the rate was initially higher for females but became more common among males. Additionally, the incidence rates within each age group increased during 1990-2000, indicating that the overall increase in incidence was not limited to specific ages.

Public health data are routinely collected and disseminated by aggregation at geopolitical boundaries. Quantification of spatial variation in incidence across counties or states

can help direct scarce public health resources for prevention to the most affected areas; however, fine-scale (i.e., within a county or "sub-county") differences in vector-borne disease risk are potentially blurred through this process (Eisen et al. 2006; Eisen and Eisen 2007; Winters et al. 2010). Understanding reasons for spatial variation in disease occurrence can help refine knowledge of risk factors for disease.

Environmental risk factors for Lyme disease

Lyme disease can occur where ticks, infected reservoir hosts, and deer coexist at a density high enough to support the dynamic enzootic cycle. Risk of acquiring Lyme disease is based on tick abundance and infection prevalence, as well as human interaction with these foci. Most epidemiologic and ecologic studies of Lyme disease have focused not on human behavior, but on elucidation of environmental risk factors—non-human factors in the environment that are associated with elevated tick density, infection prevalence, or human disease incidence.

Blacklegged tick density and infection prevalence vary across space but specific reasons for this heterogeneity are not well understood (Wilson 1998). The most likely explanations are variation in habitat type and host population movement. The highest densities of host-seeking *I. scapularis* occur in forested environments (Piesman and Spielman 1979; Maupin et al. 1991; Stafford and Magnarelli 1993; Ostfeld et al. 1995). Spatial autocorrelation of blacklegged tick abundance has been shown at state and regional scales (Nicholson and Mather 1996; Kitron and Kazmierczak 1997). In a more fine-scale analysis in South Kingstown, Rhode Island, distribution of ticks did not have any spatial pattern, but was very heterogeneous even between locations < 200 m apart (Pardanani and Mather 2004). Several studies have examined

landscape or climatic associations with *I. scapularis* abundance but without examination of human disease occurrence (Maupin et al. 1991; Stafford and Magnarelli 1993; Dister et al. 1997; Guerra et al. 2002; Allan et al. 2003; Bunnell et al. 2003; Lubelczyk et al. 2004; Diuk-Wasser et al. 2012). In another study in Rhode Island, although tick abundance and infection prevalence tended to be higher at residences of patients with reported Lyme disease than those without, those measures alone did not predict human disease (Connally et al. 2006). Because entomologic measures alone are not perfect correlates for human disease risk, human disease or tick encounters are critical study endpoints when identifying risk factors.

Consensus on established environmental risk factors for Lyme disease is lacking, possibly due to the use of varying spatial scales in analyses and continuing evolution of spatial analytic methods (Connally et al. 2006; Killilea et al. 2008). The spatial scale of analysis is critical when examining potential Lyme disease risk factors. Use of coarse geographic boundaries as entities in defining level of disease risk is subject to the loss of individual-level information in all "ecologic" analyses, and subsequent ecologic fallacy (Rothman et al. 2008). Ecologic fallacy occurs when aggregate factors tied to areas with higher disease incidence are incorrectly interpreted as individual-level risk factors. These types of studies can be valuable for hypothesis generation but cannot be used to understand human risk of Lyme disease at smaller scales (Morgenstern 2008; Rothman et al. 2008).

Rural or low-density residential development has been positively associated with Lyme disease risk when compared to medium- and high-density development areas (Glass et al. 1995; Cromley et al. 1998; Orloski et al. 1998; Smith et al. 2001). Several studies have also identified

the presence or close proximity of forests as a primary risk factor for Lyme disease at various spatial scales (Maupin et al. 1991; Glass et al. 1995; Dister et al. 1997; Kitron and Kazmierczak 1997; Jackson et al. 2006a).

Several researchers have demonstrated that, in addition to the mere presence of forest, fragmented forest characteristic of modern suburban and exurban development is associated with higher entomologic indices or human disease incidence (Allan et al. 2003; LoGiudice et al. 2003; Jackson et al. 2006a; Jackson et al. 2006b; Keesing et al. 2009). However, others found no association or an inverse association (Cromley et al. 1998; Brownstein et al. 2005; Diuk-Wasser et al. 2012). The potential association between forest fragmentation and human Lyme disease is complex. Landscape features affect animal populations; some researchers theorize that landscape fragmentation creates local foci of elevated *B. burgdorferi* prevalence by decreasing the diversity of small mammals that can live in that environment (LoGiudice et al. 2008; Keesing et al. 2009). The fragmented ecosystem favors the predominant infectious reservoir host, the white-footed mouse, which can flourish even in very resource-restrictive environments. Furthermore, forest edges resulting from fragmentation are the preferred habitat of white-tailed deer, the presence of which supports elevated tick reproduction in those areas. Finally, land-use changes that favor fragmented reforestation and suburbanization have increased human exposure to forested environments such as those bordering lawns, golf courses, or parks (Patz et al. 2004; Jackson et al. 2006a).

Several case-control studies have examined individual-level risk factors for human Lyme disease in the United States. An age-matched case-control study including data from 51 cases in

Hunterdon County, New Jersey, identified rural residence (odds ratio [OR]: 14.0, 95% confidence interval [CI]: 1.7-116.4), performing brush clearing activities (OR: 4.0, 95% CI: 1.4-14.4), and presence of a bird feeder on property (OR: 3.2, 95% CI: 1.0-10.2) to be associated with human Lyme disease (Orloski et al. 1998). Similarly, an unmatched case-control study using 294 incident Lyme disease cases in Chester County, Pennsylvania, found rural residence (OR: 2.96, 95% CI: 1.22-7.18), homes within 100 feet of forest (OR: 4.26, 95% CI: 1.71-10.59), and deer observed on the property (OR: 2.71, 95% CI: 1.94-3.79) to be risk factors for infection (Smith et al. 2001). Connally et al. found woods adjacent to property to have an elevated, but not statistically significant, effect on risk of Lyme disease in a large age- and neighborhood-matched case-control study of 349 Lyme disease cases conducted in Connecticut (OR: 1.53, 95% CI: 0.79-2.95)(Connally et al. 2009). In an age- and sex-matched case-control study of risk factors for Lyme disease in children, Klein et al. (1996) found insufficient evidence of associations between Lyme disease risk and presence of animals (including deer) on property, time spent outdoors, use of insect repellent and tick checks, along with many other potential risk factors among 44 case-control pairs (Klein et al. 1996). Using geocoded residence of reported Lyme disease patients and randomly selected control addresses in Baltimore County, Maryland, Glass et al. (1995) identified several environmental risk factors for Lyme disease: residence in forested areas, non-highly developed areas, specific soil types, and residence in two specific regions of the County (Glass et al. 1995).

Deer density and ticks on deer as environmental risk factors for Lyme disease

Brief history of white-tailed deer in the northeastern United States

Although present in moderate numbers during colonial times, deer became scarce by 1900 due to habitat loss for agriculture and unregulated hunting (Northeast Deer Technical Committee 2009). Following gradual reforestation and additional loss of predators, deer populations grew exponentially. White-tailed deer prefer habitat where forested and developed landscapes meet, an increasingly common setting as land use patterns continue to favor fragmentation in the northeastern United States (Northeast Deer Technical Committee 2009). The deer population of Connecticut increased from 19 in 1896 to an estimate of over 76,000 in 2004 (Stafford 2007). Conflicts between humans and deer have become a prominent wildlife management concern and center around three major issues: risk of zoonotic diseases including Lyme disease, deer-motor vehicle collisions, and ecologic and landscape damage (Conover 2002). Communities throughout the northeastern United States have instituted active deer management programs due to widespread overpopulation (Northeast Deer Technical Committee 2009).

Role of deer in the enzootic cycle

Deer are a preferred host for adult *I. scapularis*; they serve as sources of blood as well as a mating location (Wilson et al. 1990b). Upon completion of the blood meal, adult female ticks drop off of deer and lay eggs. Eggs hatch as larvae, seek and ingest a blood meal on small rodents, overwinter, and emerge as nymphs the following spring. As the nymphal stage blacklegged tick is responsible for most human Lyme disease illness in the eastern United States, the role of deer in human risk is separated by two years of tick life cycle and several

other factors that affect tick survival and infection prevalence (Figure 1.7.). Meteorologic factors and availability of blood meal hosts affect the survival of eggs to larvae and larvae to nymphs (McEnroe 1977; Carey et al. 1980; Carey et al. 1981; Main et al. 1981; Main et al. 1982).

The population explosion of white-tailed deer has been implicated in the emergence of several zoonotic tickborne diseases in the Northeast U.S. (Steere 2001; Telford III 2002; Stafford 2007). On the contrary, some have suggested this correlation is a coincidence given the large populations of white-tailed deer elsewhere in the United States and the possibility that small mammal hosts are more important in the enzootic cycle (Ostfeld 2011; Levi et al. 2012). Nevertheless, deer population reduction is often cited as a possible control measure for Lyme disease (Hayes and Piesman 2003; Stafford 2007).

The function of white-tailed deer in the maintenance and rate of infection within *B. burgdorferi* enzootic foci is not well understood. Although deer are a seemingly necessary factor for survival of newly introduced blacklegged ticks, once established in an area the relationship between deer and blacklegged ticks may change if alternate blood meal hosts are available (Levi et al. 2012).

Deer density

Evidence that deer population reduction corresponds to Lyme disease risk reduction is limited. Multiple observational or deer exclusion studies have been conducted; although most have shown some level of correlation between deer abundance and tick abundance, others have found no association (Piesman et al. 1979; Anderson and Magnarelli 1980; Schulze et al. 1984; Wilson et al. 1985; Wilson et al. 1990a; Daniels et al. 1993b; Stafford 1993; Duffy et al.

1994; Daniels and Fish 1995; Ginsberg and Zhioua 1999; Talleklint-Eisen and Lane 2000; Schulze et al. 2001a; Rand et al. 2003; Ginsberg et al. 2004; Jordan and Schulze 2005). Experimental studies have demonstrated that complete or near complete elimination of deer in isolated settings may have a substantial impact on reproduction of blacklegged ticks (Wilson et al. 1984; Wilson et al. 1988; Rand et al. 2004). More moderate deer reduction efforts appear to have a non-uniform effect on tick abundance, but methodological differences prevent direct comparison between studies (Deblinger et al. 1993; Stafford et al. 2003; Jordan et al. 2007). Only two studies assessed the impact of incremental deer reduction on human Lyme disease occurrence. Both studies failed to demonstrate an impact on human risk, although they were hampered by unreliable methods of disease reporting (Kilpatrick and LaBonte 2003; Jordan et al. 2007).

Deer as moderators of local infection prevalence

Studies have demonstrated conflicting information regarding the relationship between deer density and *B. burgdorferi* infection prevalence in ticks (Amerasinghe et al. 1992; Pichon et al. 1999; Rand et al. 2004). Deer also provide blood meals to immature ticks, and because they are not reservoirs, may serve to limit the local infection prevalence (Lacombe et al. 1993). Removal of deer could additionally serve to increase the number of adults seeking blood meals, increasing the likelihood of human contact with infected adult ticks (Rand et al. 2004). Deer population reduction could have unintended ecologic consequences, and the full range of impact of deer reduction on the enzootic cycle needs further research.

Tick abundance on deer

Tick population density has been compared across space and time by examining the abundance of ticks on hosts, including deer (Amerasinghe et al. 1992; French et al. 1992; Amerasinghe et al. 1993). The quantitative relationship between deer density and blacklegged tick abundance on deer is not well understood (Deblinger et al. 1993; Glass et al. 1994; Rand et al. 2003). One possible reason for discrepant findings is the sex ratio of deer. Several studies have indicated that male deer have higher levels of tick infestation than female deer, possibly because they cover a broader geographic range (Main et al. 1981; French et al. 1992; Kitron et al. 1992; Amerasinghe et al. 1993; Cortinas and Kitron 2006). Thus, the complex relationship between deer density and tick abundance and Lyme disease risk may also depend on the sex ratio of deer herds. Deer management in the Northeast United States is geared toward reduction in female deer to assist population management (Northeast Deer Technical Committee 2009); however, male deer may be more important to the Lyme disease enzootic cycle.

Blacklegged tick abundance on deer has been studied as a proxy outcome measure for human Lyme disease risk. In Indiana, clustering of blacklegged ticks on deer in space and time was detected using a spatial scan statistic (Keefe et al. 2009). Abundance on deer has been associated with environmental factors in geographic information system (GIS)-based analyses in studies in Maryland, Illinois and New Jersey (Schulze et al. 1984; Kitron et al. 1992; Glass et al. 1994). In Illinois, abundance of ticks on deer was highest in forested areas with sandy soils, near streams. In Kent County, Maryland, tick abundance on deer was associated with well-drained sandy soils, and negatively associated with urban land use and wetlands (Glass et al. 1994). In

New Jersey, most of the variability of abundance of ticks was due to elevation (Schulze et al. 1984).

Abundance of blacklegged ticks on deer has been associated with human Lyme disease at the county level and between communities > 20 km apart (Wallis et al. 1978; Schulze et al. 1984; Daniels et al. 1993a; Kitron and Kazmierczak 1997; Raizman et al. 2012). One recent study was the first to examine the spatial association between deer density at the county level and *B. burgdorferi* infected ticks on deer with human disease; however, the study was conducted in Indiana, an area with very low enzootic transmission of *B. burgdorferi* (Raizman et al. 2012). The potential association of tick abundance on deer and sub-county variation in deer density has not been evaluated.

Summary of deer density and tick abundance on deer

Additional observational or experimental studies are needed in order to determine the potential for localized reduction in deer density to serve as a viable Lyme disease prevention measure. The association between sub-county spatial variation in deer density and human disease has not been examined. Deer populations fluctuate dynamically, but deer generally have a limited home range of up to one mi^2 (Sparrowe and Springer 1970; Kilpatrick and Spohr 2000; New Jersey Department of Environmental Protection 2010). In a suburban landscape, deer home range is smaller than in forested or agricultural environments. One study in Connecticut using telemetry data revealed a mean home range of 0.17 mi^2 for female white-tailed deer that did not vary significantly by season (Kilpatrick and Spohr 2000). Small area variations in deer density may be partially responsible for fine-scale variation in distribution of

blacklegged ticks in the environment. The biologic mechanism through which some previously described environmental risk factors (such as density of development or forest fragmentation) may be associated with Lyme disease risk may be through fine-scale differences in deer density. Evaluation of the associations between static environmental risk factors, deer density, tick abundance on deer and human Lyme disease is needed in order to better understand causes of spatial variation in disease risk.

FIGURES

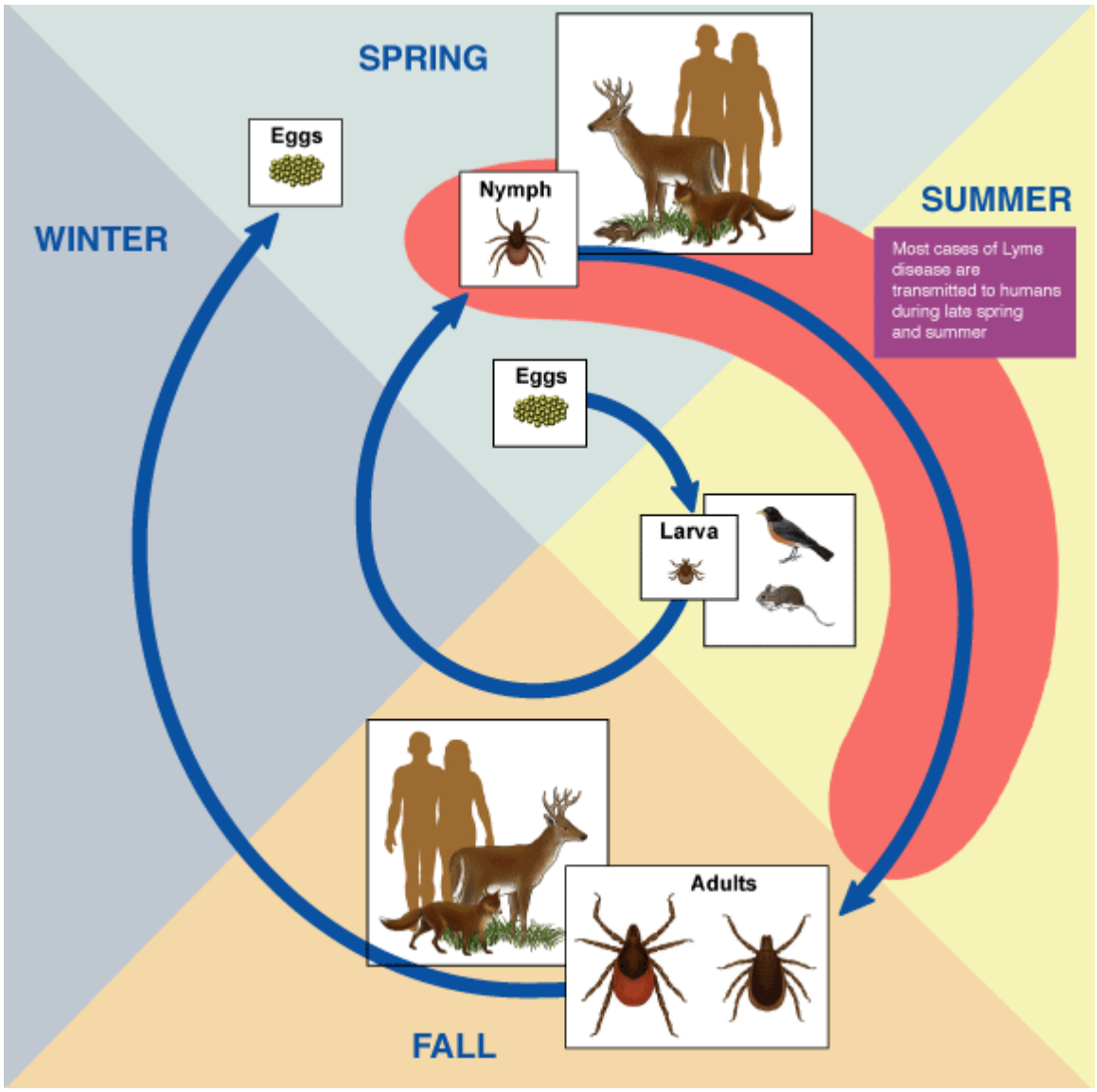


Figure 1.1. Illustration depicting the enzootic cycle of Lyme disease

(Courtesy: CDC-Division of Vector-Borne Diseases)

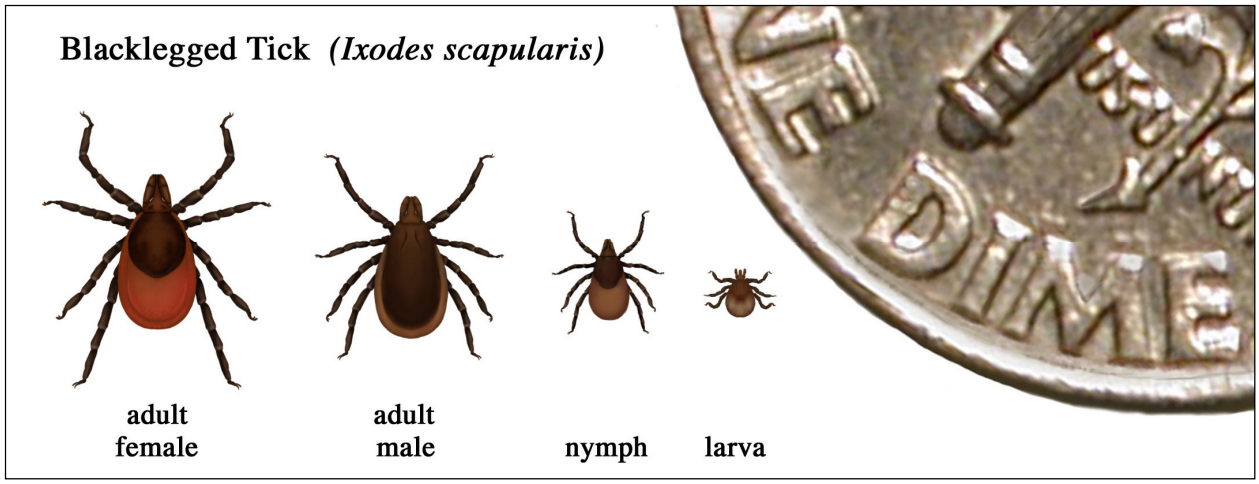


Figure 1.2. Illustration of the life stages of *Ixodes scapularis* ticks

(Courtesy: CDC-Division of Vector-Borne Diseases)

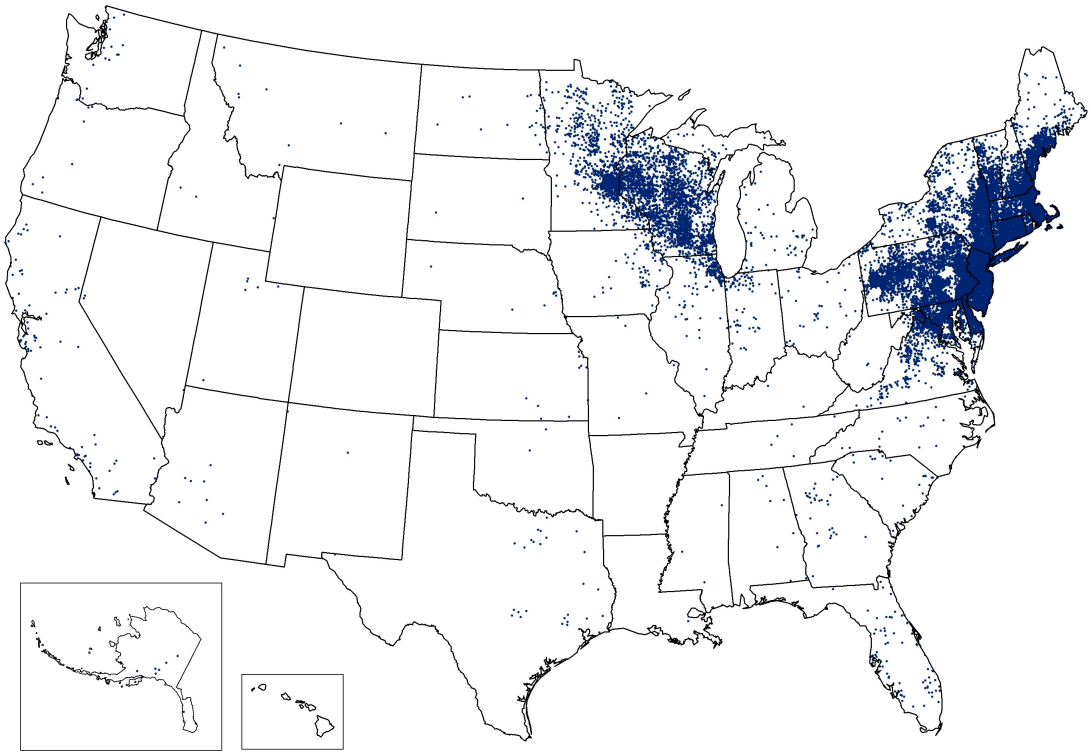


Figure 1.3. Confirmed Lyme disease cases, United States, 2011

(Courtesy: CDC-Division of Vector-Borne Diseases)

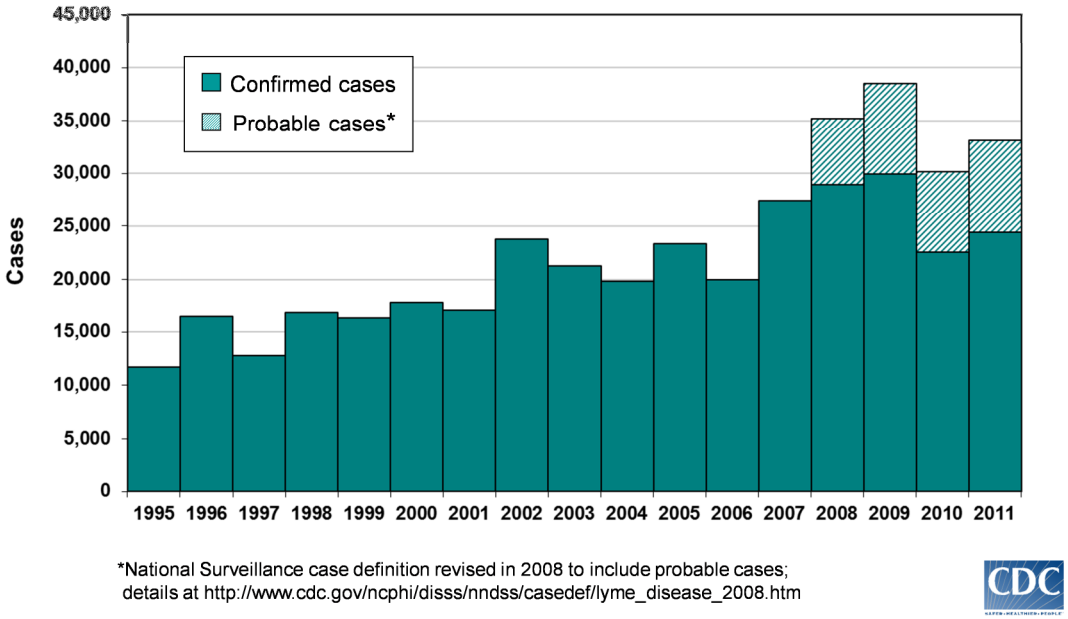


Figure 1.4. Number of reported Lyme disease cases, by year, United States—1995-2011

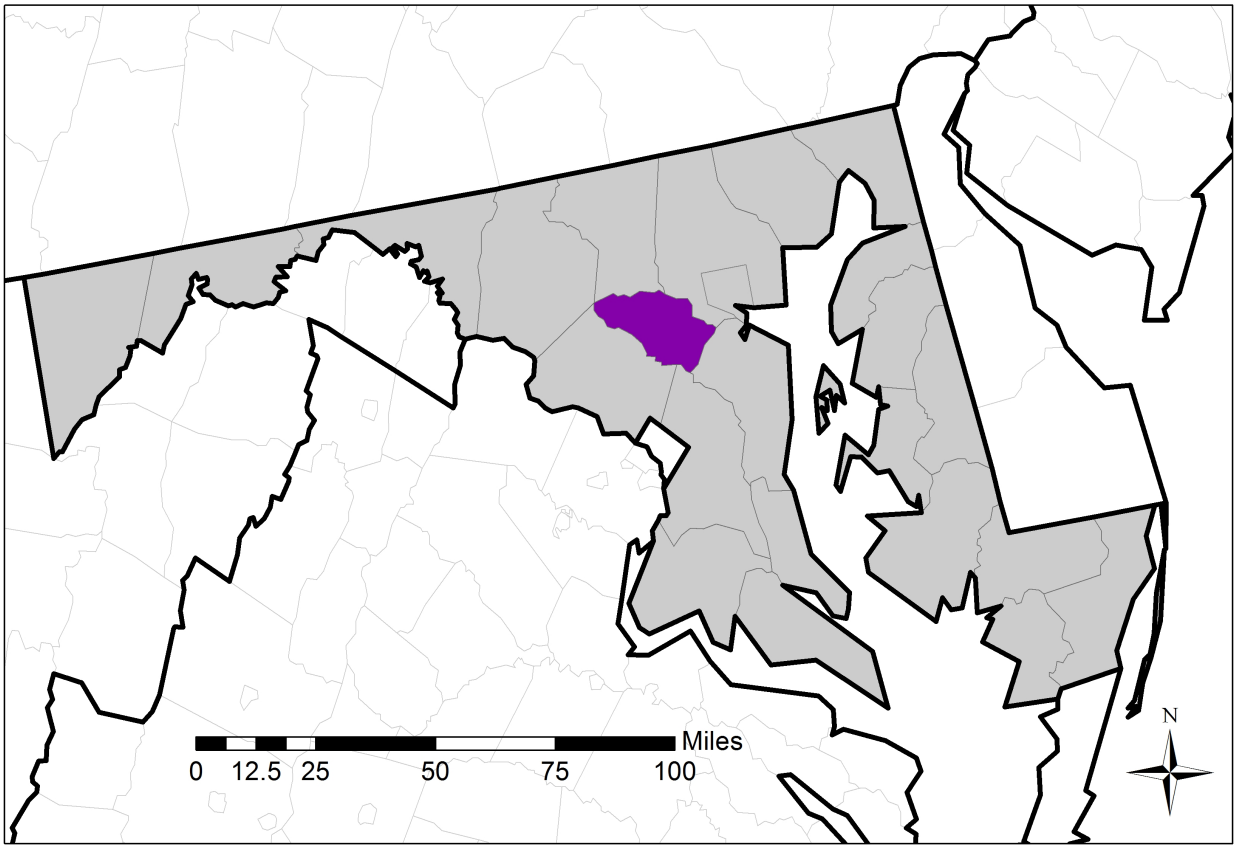


Figure 1.5. Map of the State of Maryland highlighting Howard County

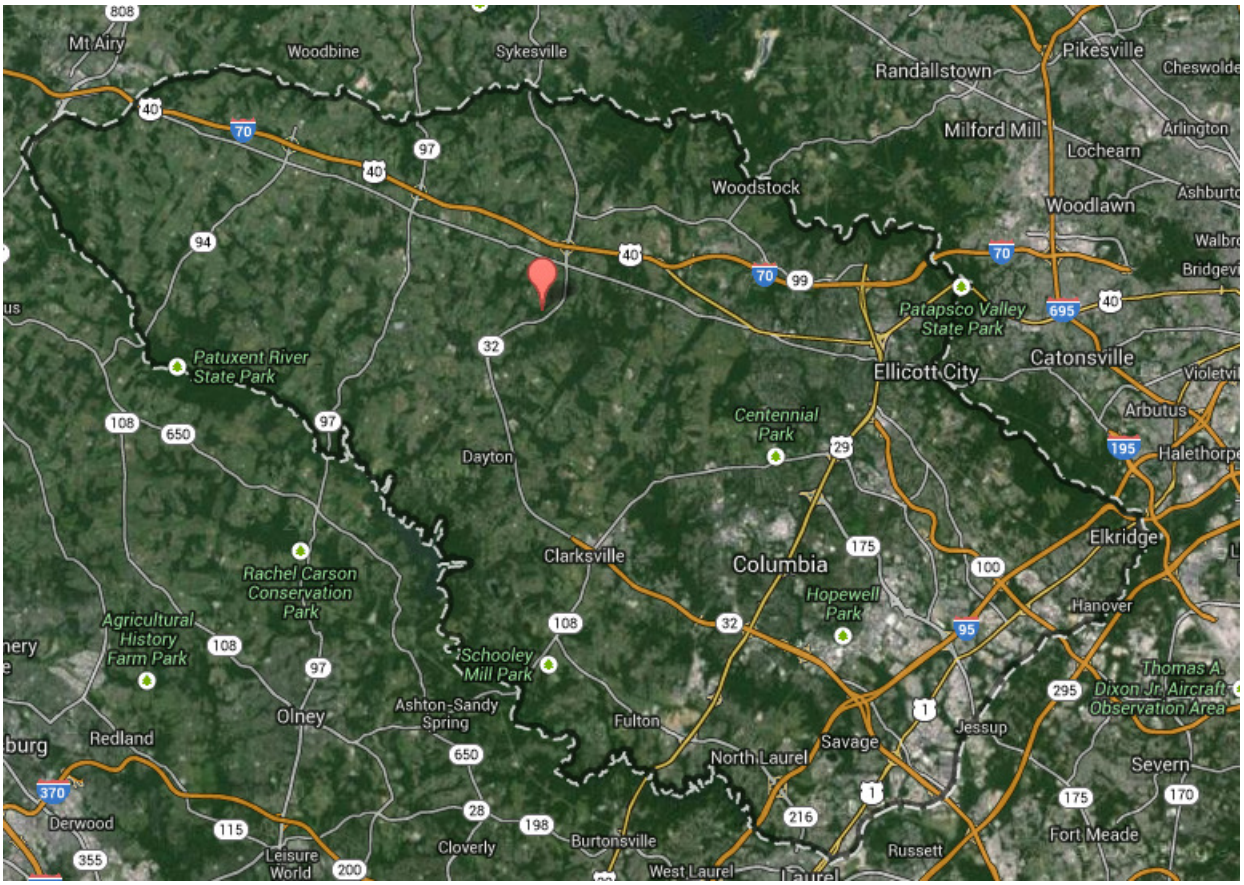


Figure 1.6. Satellite image of Howard County, Maryland

(source: www.maps.google.com)

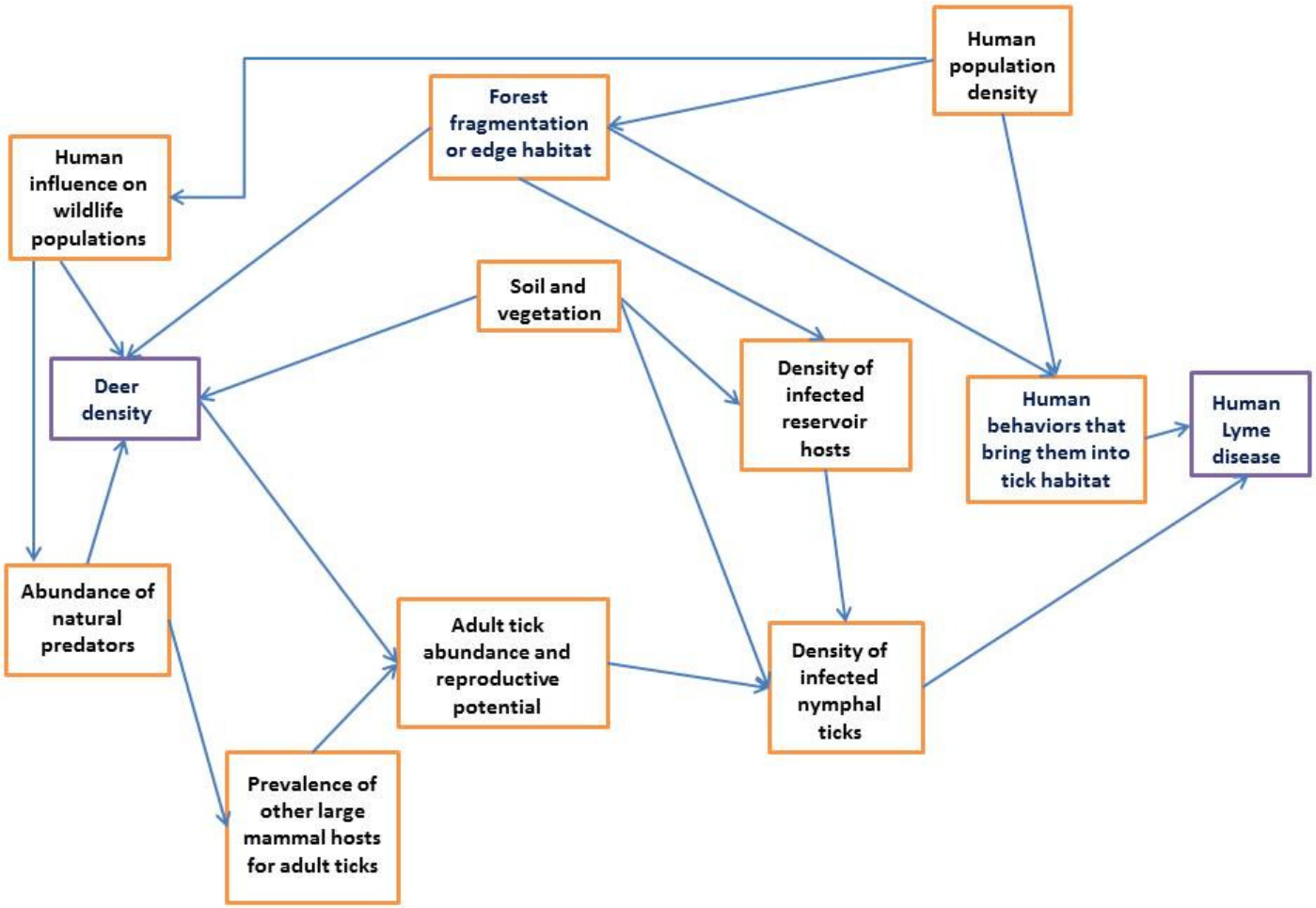


Figure 1.7. Pathway between white-tailed deer, other environmental factors, and human Lyme disease*

*Deer density and human Lyme disease are in purple boxes while other related or intermediate factors are displayed in orange boxes

CHAPTER 2: PROJECT 1

2. EPIDEMIOLOGY AND CLUSTERING OF HUMAN LYME DISEASE, HOWARD COUNTY, MARYLAND—2001-2011

INTRODUCTION

In the United States, Lyme disease incidence and geographic range have increased since the disease was first described in the late 1970s, despite several possible prevention methods (Hayes and Piesman 2003; Bacon et al. 2008). Lyme disease occurs across broad geographic areas; however, spatial and temporal variation in blacklegged tick abundance, tick infection prevalence, and human disease has been documented within these areas. Variation in abundance of blacklegged ticks in the environment has been documented not only between states and regions, but also between locations only a few hundred meters apart (Pardanani and Mather 2004; Killilea et al. 2008).

Evaluation of spatial and temporal variation in human disease has been relatively infrequent (Glavanakov et al. 2001; Frank et al. 2002; Chen et al. 2005; Chen et al. 2006; Waller et al. 2007). Comparing human Lyme disease incidence across states and counties and over time quantifies movement and trends as the causative agent moves into new areas, although it does not provide fine-scale information on spatial variation in risk within those counties. Examining sub-county variation in disease occurrence may provide critical information regarding disease distribution and possible risk factors. Description of clustering of human Lyme

disease cases within a single county can provide insight into the fine-scale processes that result in human disease within that county and refine hypotheses regarding risk factors for disease. However, the spatial or spatiotemporal clustering of human Lyme disease has not been quantified on a sub-county scale. Furthermore, cluster size itself may reveal a spatial scale at which control measures and education efforts are most appropriate (Glavanakov et al. 2001).

Recent developments in software that utilize spatial scan statistics allow for statistically robust determination of areas with elevated risk of disease. The goals of this project were to describe the epidemiology and clustering of human Lyme disease in one county and to explore environmental and socio-demographic factors that may be associated with areas of elevated disease risk.

METHODS

Study location

The State of Maryland consistently reports a high incidence of human Lyme disease (Bacon et al. 2008; Centers for Disease Control and Prevention 2013a). Howard County, Maryland, is located between Baltimore and Washington D.C. Its 250 square miles had a 2010 population of approximately 287,000 people, with a density of 1,145 people per square mile (US Census Bureau 2012). Howard County was selected as the study location because of the emergence of Lyme disease during the study time frame and availability of detailed geographic information.

Study population

Potential Lyme disease cases that occurred during 2001-2011 were reported by clinicians and laboratories, and subsequently investigated and classified according to standardized case definitions by local and state health officials. Cases occurring 2001-2007 were classified according to the case definition established in 1996 (Council of State and Territorial Epidemiologists 1996); cases occurring 2008-2011 were classified according a revised case definition that took effect in 2008 (Council of State and Territorial Epidemiologists 2008). Data submitted from the Maryland Department of Health and Mental Hygiene (MDHMH) to the Centers for Disease Control and Prevention (CDC) through the National Notifiable Diseases Surveillance System (NNDSS) were used for basic descriptive analyses; these data included age, sex, date of disease onset, presence of symptoms associated with the surveillance case definition, and case status (i.e., confirmed or probable). Additionally, for the sole purpose of this study, MDHMH directly provided Lyme disease case data that were stripped of patient age and sex information, but contained residential address (not transmitted through NNDSS). Methods for this project were approved by Institutional Review Boards at CDC, MDHMH, and Colorado State University (Appendix 1.0.).

Geographic data

Addresses for all residential structures as of 2004 were available from the Howard County Geographic Information Systems Division (n=94,308). Households present in this file were assumed to be all households present during the entire study period, 2001-2011. Residential address data from reported Lyme disease patients were cleaned for misspellings and typographical errors with Google Maps and geocoded using the Howard County geographic

household file within ArcGIS v.10.1 (Environmental Systems Research Institute [ESRI], Redlands, CA). Census block group data (including year 2000 block group boundaries, 2009 population estimates, 2009 estimates for educational, age distribution, income and home value variables, and 2000-2009 population growth estimates) were obtained from ESRI.. A 2010 land use dataset was obtained from the State of Maryland Department of Planning (resolution: one meter). A 2002 land cover dataset was obtained from the United States Geological Survey Gap Analysis Program (resolution: 30 meters). Metrics of forest fragmentation were calculated using Patch Analyst, the FRAGSTATS™ software add-in to ArcGIS (Rempel et al. 2012).

Analytic methods

Mid-year 2005 U.S. Census Bureau estimates were used to calculate Lyme disease incidence in Howard County, neighboring counties, and the State of Maryland.

Case density surface

A visual assessment of the spatial variation in Lyme disease by household during the study period was accomplished by creating a smooth density surface within ArcGIS. Kernel density functions, which interpolate point data into a continuous surface, were used to create smoothed surfaces. Two smoothed surfaces were created: one for case households and one for the population denominator of all households within the County. Kernel density functions were calculated using a bandwidth of 2640 feet (1/2 mile), with cell sizes of 370 feet. A case density surface (analogous to a map of cumulative disease incidence) was calculated by dividing the two surfaces (Peterson et al. 2009). The resulting map displayed cumulative incidence of case households per total Howard County households.

Cluster detection

Home address was assumed to reflect the spatial location of exposure; cluster detection was conducted using the spatial scan statistic available in SaTScan™ software v.9.1.1 (Kulldorff M. and Information Management Services Inc. 2011). The spatial scanning software examined circles of varying radii and calculated a relative risk (RR) for each evaluated cluster based on the observed number of households with reported Lyme disease within the circle versus the number expected under the null hypothesis. The null hypothesis was that disease risk was the same within any evaluated circle as it was outside of the circle (Kulldorff M. and Information Management Services Inc. 2011). Cluster analysis utilized the geographic coordinates of each household in Howard County. Households were classified as having at least one reported case of Lyme disease (case households) or no reported Lyme disease during the entire study period (non-case households). Spatial cluster detection calculations were based on a binomial distribution.

For spatiotemporal cluster detection, the method was similar to that of spatial cluster detection, except the scan assessed cylinders of various sizes and heights, where the size was equivalent to the circle evaluated in a spatial scan and the height of the cylinder was equivalent to time. For this type of analysis, census block group was used as the administrative boundary and assessment of potential clusters was based on a Poisson distribution of case counts occurring within census block groups. Based on geographic location, case households were assigned to one of 118 census block groups located within Howard County from the year 2000 Census. Reported date of disease onset was used to represent time of infection; analyses scanned windows with one-month intervals. If more than one case occurred at the same

address, additional cases were included in spatiotemporal calculations that did not factor into spatial analyses (based on binary outcome of at least one case per household vs. no cases).

Tests of statistical significance for each potential spatial or spatiotemporal cluster were based on a likelihood ratio test. Monte Carlo simulations (n=999) were used to obtain p-values as per the software default methodology. A high-risk disease cluster was defined as an area with an elevated RR and a p-value < 0.05. Analyses were conducted using the software default maximum cluster size (50% of the study population) and at several different maximum cluster sizes in order to better resolve the shape and distribution of clusters according to residence. Confirmed and probable Lyme disease cases were used. Clusters were visualized and additional spatial statistics calculated in ArcGIS.

Exploratory description of factors associated with high-risk clusters

To describe how geographic areas identified as high-risk clusters differed from remaining areas of the County, environmental and socio-demographic characteristics were compared between these areas. Environmental factors from land use and land cover datasets were extracted as total area units inside and outside high-risk clusters using the Tabulate Area tool in ArcGIS. Cell counts for individual land use and land cover categories were combined into broad categories for analysis. Specifically, land use categories were the following: very low-density development dominated by agriculture or open fields (properties between five and 20 acres), very low-density development dominated by forest (properties between five and 20 acres), low-density development (areas of at least 90% single family or duplex dwellings, with lot sizes of from one-half acre to five acres), medium-density development (lot sizes between

one-eighth and one-half acre), high-density or urban development (more than eight dwelling units per acre or other urban non-residential uses), herbaceous or agricultural use (e.g., parks, pastures, cropland, and brush), and forest (deciduous, evergreen, or mixed forest types), herbaceous and agricultural (pasture, row crops, parks). The final land cover categories were the following: urban (impervious surfaces), herbaceous or agricultural (i.e., row crops, cultivated trees, pasture), red oak and white oak forest (oak-hickory forests with forest canopy dominated by *Quercus* [*alba, rubra*] and other oaks), chestnut oak forest (dominated by *Quercus prinus* with other oaks [*Q. alba, falcata, rubra, velutina coccinea*]), all other deciduous forest classes, and mixed deciduous/evergreen forest.

Forest fragmentation metrics were calculated using the forest class of the land use dataset. Specifically, amount of forest edge relative to total forest area and mean forest patch size were calculated. Census variables such as median income and educational attainment were calculated as area-weighted averages of census block group values for both geographic areas (inside high-risk clusters and outside clusters).

Proportion of total area inside high-risk clusters in each land use or land cover category was compared to proportion of total area similarly classified in the remaining area of the county using χ^2 tests in SAS (SAS Institute, Cary, NC). Due to the nature of the total area data extraction and modifiable areal unit problem that plagues area-based geographic analyses, statistical tests were conducted at two different units of area. Specifically, area was converted to units of one hectare (ha) (i.e., 0.01 km^2) and one km^2 for statistical tests, and corresponding results compared. In contrast to the statistical testing for comparison of land use and land cover

between high-risk clusters and the remainder of the County, forest fragmentation and socio-demographics were described in the respective areas but without accompanying statistical tests. Given the area-weighted calculations for these variables, design of advanced statistical methods to compare these calculations was beyond the scope of our exploratory objective.

RESULTS

Incidence

From the NNDSS database, 1,914 confirmed and probable Lyme disease cases were reported from Howard County during 2001-2011. Annual incidence in Howard County ranged from 14.5 cases to 132.1 cases per 100,000 residents during the study period (mean: 64.6 cases per 100,000 residents; Table 2.1). During 2001-2011, Lyme disease incidence in Howard County varied from roughly equal to three-fold higher than the incidence for all of Maryland (Figure 2.1.). Despite year-to-year variation as would be expected due to underlying variation in the enzootic cycle and surveillance practices, incidence in most Maryland counties increased during the beginning of the study period, peaking in 2007 or 2008 (Figure 2.1.).

Characteristics of reported Lyme disease patients

The age distribution of reported cases among Howard County residents was bimodal; cases occurred most commonly among 5-14 year olds and 40-65 year olds. The median age of cases was 42 years (range: < 1 – 89 years; Table 2.1.). The age distribution did not change substantially from year to year; the lowest median ages of 28 and 28.5 occurred in the years with the fewest number of reported cases (2002 and 2003). Males accounted for 56% of cases; the male predominance occurred in all years except that with the fewest total cases (2002; Table 2.1.). Race and ethnicity were unavailable for the majority of reported cases. Among the

1,303 confirmed cases with at least one confirmatory clinical sign or symptom reported, erythema migrans (EM) rash was the most common clinical sign (76%), followed by arthritis (12%), Bell's palsy (12%), radiculoneuropathy (7%), meningitis (2%), encephalitis (1%), and atrioventricular block (1%). The percent of confirmed cases with patient symptom information that had reported EM rash varied over the study period, ranging from 54% in 2005 to 91% in 2009. Among all confirmed cases, the percent of patients with EM ranged from 53% in 2005 to 74% in 2011. Five percent of cases were hospitalized. The median week of disease onset was 26 (approximately end of June), but varied between week 24 (mid-June) and week 29.5 (end of July; Table 2.1.). Annual incidence was not correlated with week of median disease onset (r=-0.34; p=0.306). Median week of disease onset and median age did not differ according to surveillance case status (confirmed vs. probable; Wilcoxon Rank Sum test p=0.701 and p=0.555, respectively). The general epidemiology of Lyme disease cases within Howard County thus generally mirrored that of the nationwide picture (Bacon et al. 2008). Appendix 1.1 describes surveillance practices in Howard County during the study time period and discusses possible effects of these practices on study findings. Incidence of EM cases alone, those with a disease manifestation likely less subject than others to changes in surveillance practices, demonstrated a similar temporal trend when compared to the trend in overall incidence. This finding suggested that temporal changes in reported case numbers during the study period likely reflected true changes in disease occurrence rather than surveillance artifact (Appendix 1.2.).

Geographic information

Residential address of patients with reported Lyme disease during 2001-2011, devoid of other personal identifiers, was provided directly to CDC from MDHMH. The total number of

cases (n=1,934) differed slightly from that reported through NNDSS, likely because of differences in timing of when reporting years are closed to modification (K. Feldman, MDHMH, personal communication). Of these, 112 cases (6%) were excluded from further geographic analyses because they lacked an address or only listed a post office box and 52 (3%) were excluded because listed addresses were not located within Howard County. In total, 1,753 (91%) cases (1,521 confirmed, 232 probable) from 1,672 unique residences were successfully geocoded (Table 2.2.). Among the 80 cases that occurred in the same household as another case, some may have been duplicate entries of the same case (e.g., the same onset date, but a repeated positive serology the following year); duplicate entries could not be reconciled definitively given lack of identifiers.

The density surface of human Lyme disease according to residence displayed a heterogeneous, patchwork pattern of disease occurrence across the County (Figure 2.2.). Density of human Lyme disease was lower in the eastern part of the County compared to the central and western parts of the County. Examination of cumulative incidence across census block groups revealed a similar pattern (Appendix 1.3.).

Cluster detection

Spatial cluster analyses

Unique geocoded case households (n=1,672) and non-case households (n=92,636) were input into SaTScan software for spatial cluster analyses. Analyses were conducted at several maximum cluster sizes (Figure 2.3.). Geographic clustering of Lyme disease cases by residence was evident, regardless of maximum cluster size setting within SaTScan software. Overall, high-

risk clusters were in the west-central part of the County. With the software default maximum cluster size of 50% of the entire population, nearly all of the western and central part of Howard County was part of one very large high-risk cluster with a 15 km radius. The corresponding RR of 2.0 indicated that, in this area, risk of a household having reported Lyme disease was twice that of the rest of the County; two additional smaller high-risk clusters were also identified (Figure 2.4., top panel). The greatest number of high-risk clusters (n=7) were detected when setting the maximum cluster size to two km radius (Figure 2.3.). This setting also detected high-risk clusters that were more specific to geographic areas with high case density (Figure 2.4., bottom panel). High-risk clusters identified at the two km radius maximum displayed RRs that ranged from 2.1-12.9 (Table 2.3.). The potential for population growth during the study period to influence cluster detection due to artificial underestimation of the population at risk was minimal (Appendix 1.4.). Using case counts per census block group, rather than point process data, clusters detected with the software default maximum size were similar (Appendix 1.5.).

Spatiotemporal cluster analysis

Spatiotemporal cluster detection was based on census block group population denominators. The average 2009 estimated population per census block group in Howard County was 2,443 (range: 571-7,371 people). The mean number of reported Lyme disease cases per census block group was 15 (range: 0 - 68 cases). Spatiotemporal analysis conducted with the software default maximum cluster size revealed a very large cluster located in virtually the same spatial location as detected in the spatial cluster analysis, encompassing the entire western and central part of the County during 2005-2011 (Appendix 1.6.).

Description of factors associated with high-risk clusters

Exploratory analyses were conducted to compare environmental and socio-demographic factors inside high-risk clusters using the two km maximum radius setting to the remaining area in the County. Due to the sheer volume of area-based pixels extracted as each land use and land cover class, nearly all factors were statistically significantly different inside high-risk cluster areas when compared to the remaining area using the one ha unit of analysis, despite very minimal absolute difference (Table 2.4.). Comparison using a larger unit (one km^2) of analysis (and consequently decreasing the total N [or sample size] in the statistical test) revealed potentially more biologically-meaningful significant differences. Regardless of the unit of analysis, when compared to area outside, high-risk clusters had more low-density residential development (30.3% vs. 16.7% outside of clusters) and more red and white oak forest (38.9% vs. 15.3% outside of clusters; Table 2.4.). High-risk clusters also displayed less herbaceous cover and land used for agriculture (Table 2.4.). The proportion of land use classified as forest (and thus presumed to more likely represent contiguous forest without other uses) did not greatly differ between the areas inside and outside high-risk clusters. However, upon summation of the land cover classes that pertain to forest (likely less contiguous forest), higher amount of total forest cover was present inside the two km high-risk clusters as compared to the remainder of the County (Appendix 1.7.). Proportions of some land use and land cover classes inside high-risk clusters and outside were sensitive to maximum cluster size detection used; nevertheless, for most land use and land cover classes, findings were comparable regardless of cluster detection software setting (Appendix 1.7.).

Although interpretation of comparisons is hampered by lack of statistical testing, two separate indices suggested more fragmented forest was present within the two km maximum high-risk clusters compared to the rest of the County. Specifically, in a ratio of length of forest edge to total forest area, more forest edge habitat existed inside high-risk clusters (ratio of 68:1 vs. ratio of 20:1). Mean forest patch size inside clusters was smaller than in the rest of the County (39 acres vs. 474 acres); smaller forest patches are considered an indicator of increased forest fragmentation (Patz et al. 2004; Brownstein et al. 2005). Additionally, fragmentation metrics were calculated using land cover classified as red and white oak forest. These findings were opposite in direction to those of land use classified as forest: the ratio of red and white oak forest edge to red and white oak area was smaller inside high-risk clusters (ratio of 80:1 vs. ratio of 106:1), and average patches of red and white oak dominated forest were twice the size as those outside high-risk clusters (49 acres vs. 24 acres). The potential association between fragmentation and high-risk clusters was thus sensitive to which type of forest measure was examined.

In descriptive evaluation, the area inside the two km maximum high-risk clusters was characterized by higher median household income, higher median per capita income and higher median home value, and a higher percent of the population with at least a bachelor's degree (40% vs. 36%), despite seemingly equivalent crude age distributions (Table 2.5.). Area within high-risk clusters was also characterized by higher average population density; population growth during 2000-2009 was similar between the geographic areas (Table 2.5.).

DISCUSSION

In this first reported evaluation of the spatial variability of human Lyme disease within a single county, risk of Lyme disease varied according to geographic location of residence. In high-risk cluster areas, risk of human Lyme disease was about twice that expected based on the underlying population distribution. Clustering was detected in the same general areas of Howard County despite modification of the maximum detectable cluster size. Furthermore, high-risk clusters were robust to analysis using point process data, case count aggregates within census block groups, and spatiotemporal data. No clear spatial scale at which clustering occurred was evident, as different size high-risk clusters were identified with modified software settings.

The substantial spatial clustering of human disease was supported by noteworthy differences in land use, land cover, and degree of landscape fragmentation in these areas when compared to the remainder of the County. The geographic areas containing high-risk clusters were generally wealthier and had more low-density residential development (lots between one-half and five acres). These findings agree with results from prior studies in other locations that used different study designs, outcome measures, and analytic methods. Rural or low-density development has been associated with increased risk of human disease in studies conducted in Maryland (Glass et al. 1995), New Jersey (Orloski et al. 1998), Pennsylvania (Smith et al. 2001) and Connecticut (Cromley et al. 1998). Presence of forest has been associated previously with tick abundance and human risk (Maupin et al. 1991; Glass et al. 1995; Dister et al. 1997; Jackson et al. 2006a). Here, no difference was demonstrated in amount of contiguous forest between high-risk clusters and remaining area; however, land cover classified as forest

(representing more interspersed or fragmented forest rather than contiguous forest) was more prevalent in high-risk clusters. Higher density of forest edge habitat and smaller fragments of contiguous forest indicate that more fragmented forest may be present in the areas characterized as high-risk clusters. Nevertheless, examination of an alternate forest data source (land cover classified as red and white oak forest) provided opposite results in the same area using the same overall methodology. Indices related to forest edge and forest patch size have been associated with elevated tick abundance or elevated human incidence in other studies (Allan et al. 2003; LoGiudice et al. 2003; Jackson et al. 2006a). However, the directions of associations reported in the literature are inconsistent, as they are here.

This study was the first to describe a potential association between human Lyme disease incidence and a specific type of forest in the eastern United States. Although abundance of blacklegged ticks has been generally linked to oak and maple forests (Piesman and Spielman 1979), to our knowledge, this study is the first to link any specific forest type in this region with risk of human disease. Some forest types may provide more plentiful rodent food sources or leaf litter, or could reflect the underlying soil composition and its impact on tick survival (Jones et al. 1998b; Jones et al. 1998a; Ostfeld et al. 2001; LoGiudice et al. 2008) . Nevertheless, the finding of proportionally more red and white oak forest inside high-risk clusters could be spurious or due to its commonality as the most frequent type of deciduous forest in Howard County.

Overall, the descriptive epidemiology of Lyme disease cases from this single county mirrors the general picture of Lyme disease nationwide. Specifically, patients with reported

Lyme disease are predominantly male, the age distribution is bimodal, and EM rash is the most common clinical manifestation. Yet, year-to-year variability in these characteristics was apparent, as might be expected with relatively small number of cases in a single area.

Limitations

This study was subject to several limitations, including error introduced by surveillance practices and geocoding. Lyme disease cases were considered those that were reported to the public health system and met specific criteria of the surveillance case definition. Under-reporting is common to most surveillance systems and is also documented in Lyme disease surveillance (Coyle et al. 1996; Meek et al. 1996; Campbell et al. 1998; Naleway et al. 2002). The degree of under-reporting varies by public health jurisdiction and over time. In addition, some degree of over-reporting based on misdiagnoses is likely to occur. Inherent under-reporting or over-reporting of cases, or changes in the case definition that occurred during the study period, would result in misclassification of individual households according to Lyme disease status. The ability of the health department to detect and classify Lyme disease cases was likely not related differentially to space across the County; however, the possibility of differential under-diagnosis and under-detection in lower income areas with presumably less access to health care cannot be ruled out. This detection bias would result in a differential misclassification across space, with more reported Lyme disease cases occurring in higher income areas, and increased likelihood of detecting spatial clusters in higher income areas. Although the area inside the two km maximum clusters was characterized by higher income and home value, the remainder of the County was still, on average, far from disadvantaged, minimizing the potential for this bias to affect cluster detection. Additionally, differential detection could occur because

of differences in providers with respect to reporting practices. However, the likelihood for differential detection due to provider reporting practices tied to the spatial location of patient residence is minimized in this relatively small area, as people likely utilize clinicians that are not necessarily the closest to their house, but those clinicians that accept their insurance. Temporal trends in incidence in Howard County compared to neighboring counties and the State of Maryland were similar. This similarity suggests that the increasing and then decreasing trend in incidence over time that occurred in Howard County during the study period likely reflected true changes in disease risk rather than temporal differences in disease detection.

A total of 91% of confirmed and probable cases were able to be geocoded. Post office boxes are often associated with rural rather than suburban areas, and can bias geocoding ability away from rural areas (Zimmerman et al. 2008; Wey et al. 2009). The 38 (21%) non-geocoded cases with post office boxes for addresses were located broadly across 17 of the 35 zip codes present within Howard County. The most common post office box zip codes were located in more heavily populated areas; their mostly non-rural location suggests that the impact of this potential geographic bias likely had minimal effect on our findings. Other cases were unable to be geocoded because of missing address or irreconcilable typos in address. Misclassification of these case households as non-Lyme disease houses was likely non-differential in space across the County and thus only served to attenuate our ability to detect significant clusters.

In this study, residential address was used as a proxy for location of exposure to infected blacklegged ticks. The peridomestic exposure assumption is common in scientific literature

pertaining to Lyme disease in the eastern United States; the assumption is based on data that demonstrated presence of infected blacklegged ticks in the yards and woods surrounding the homes of Lyme disease patients and that many patients believe their property to be where they acquired a tick bite (Falco and Fish 1988; Maupin et al. 1991; Nicholson and Mather 1996; Cromley et al. 1998; Orloski et al. 1998). Peridomestic exposure is, in general, difficult to validate and could not be validated in this study. Yet, our finding of statistically significant areas of elevated risk to residents implies either: 1) most disease is acquired in the home environment or neighborhood, or 2) away from home behavior patterns are similar among neighbors (e.g., they visit the same distant parks for recreation). While some misclassification of spatial exposure location is inevitable, substantial non-differential exposure misclassification would likely limit cluster detection ability due to presence of increasedrandom error and loss of statistical power (assuming common epidemiologic understanding regarding directionality of bias applies to spatial misclassification (Rothman et al. 2008). Consequently, the finding of spatial clustering according to location of residence likely reflects some inherent environmental or social processes that put people who reside in certain areas at elevated risk for disease. For example, these areas of elevated risk may be dominated not only by specific environmental characteristics that promote increased abundance of infected ticks, but may also be characterized by more residents who interact with their environment (e.g., people who garden more or more often recreate outdoors in wooded or forest edge habitat).

The impact of areal units on statistical testing is an example of the modifiable areal unit problem (MAUP) that plagues many geographic analyses (Waller and Gotway 2004; Schabenberger and Gotway 2005). MAUP actually reflects two underlying but overlapping

issues related to ecologic fallacy. The first, the “scale effect”, occurs when different results coincide with increasing larger units of spatial areal aggregation. The second, sometimes referred to as the “grouping effect”, occurs when results differ depending on the shape in which the underlying data are aggregated. While the field of spatial statistics has yet to resolve these problems, being aware of them and understanding the limitations of the data are paramount. In this study, aggregation of areal units to two different scales revealed some differing results, but also demonstrated some consistent findings regardless of areal unit. Potentially, findings consistent across scales may be the most biologically meaningful.

This study demonstrated that Lyme disease occurrence on a sub-county level is non-uniform. Although detection of significant human disease clustering is likely a finding that is generalizable to other endemic areas, there are many reasons why the generalizability of findings of specific maximum cluster sizes and corresponding environmental and socio-demographic associations may be limited. The underlying abundance of reservoir hosts and tick infection prevalence may coalesce differentially across space in other counties and states and inherently modify human risk. The findings from this study are exploratory in nature, and demonstrate the utility of conducting cluster detection analysis in endemic areas to identify those areas at highest risk of disease. This approach can be implemented in additional geographical areas to assess generalizability of findings of cluster sizes and of associated factors.

Strengths

This study was the first to describe sub-county clustering of human Lyme disease; it was also the first study to associate environmental and socio-demographic factors with clustering of human disease. Although some of these factors have been associated with elevated tick abundance and human risk in different study designs, none have been associated with clustering of disease. Furthermore, this is the first description of an association between a single oak forest type and human disease risk in the eastern United States.

Although public health resources and policies for Lyme disease case follow up differ across counties and states, activities are often relatively consistent within a single county. Restriction of this analysis to a single county minimized potential confounding by differential case detection across space that could occur by conducting this type of study across different counties. Fortunately, this analysis was able to be conducted with point-based data for not only case households, but all households within the County. Consequently, cluster finding was not limited by the minimum size of a political unit (here, a census block group).

Use of human disease data is an important outcome measure to assess risk of Lyme disease (Cromley et al. 1998; Eisen and Eisen 2008; Eisen et al. 2012). Presumably, risk increases when the abundance of infected ticks in the environment increases. However, examination of abundance of infected ticks as an outcome measure, rather than human illness, ignores the role that human behavior plays in tick exposure. Human behavior patterns may be independent of environmental characteristics, and are difficult to measure accurately. The ultimate endpoint of human disease inherently takes some human behavior into account (Eisen et al. 2012).

Conclusions

Quantification of sub-county disease clustering and identification of factors potentially associated with that clustering are critical to understanding disease processes and to development and implementation of appropriate prevention measures. Comparing overall disease incidence across counties or states can help direct public health resources, but does not refine understanding of where risk is highest, why it may be high in specific places, and who is at highest risk for encountering infected ticks. In this study, we examined univariate differences in environmental and socio-demographic factors although many of these factors are not independent of one another. An understanding of the relative importance and potential confounding of these factors with one another in their association with human Lyme disease is needed on a sub-county scale. Project 2 builds upon this work to examine the multivariable associations of these same factors with human Lyme disease risk at the household level. Additionally, further investigation into the role of specific forest types in Lyme disease risk is needed. Conducting analyses at the sub-county level is necessary in order to better refine risk factors and understand disease processes that vary in highly endemic areas. Similar sub-county analyses should be conducted elsewhere using residential data to compare findings.

TABLES

Table 2.1. Human Lyme disease cases, Howard County, Maryland 2001-2011*

Year	**Confirmed cases** (n)	**Probable cases** (n)	**Incidence****	**Male** (%)	**Median age** (years)	**Median onset week**
2001	68	-	25.2	56	36	26
2002	39	-	14.5	44	28	26
2003	42	-	15.6	64	28.5	29.5
2004	94	-	34.9	62	43	26.5
2005	150	-	55.7	61	44.5	27
2006	113	-	41.9	55	44.5	26
2007	352	-	130.6	55	40	27
2008	296	60	132.1	52	42	26
2009	203	65	99.5	59	42	25
2010	171	66	88.0	57	43.5	24
2011	137	58	72.4	55	44	27
Total	**1,665**	**249**	**64.6**†	**56**	**42**	**26**

*Reported through the National Notifiable Diseases Surveillance System

**Total (confirmed and probable) cases per 100,000 residents, based on mid-2005 population estimate

†Average annual incidence (cumulative incidence=710.3 cases per 100,000 residents)

Table 2.2. Proportion of reported Lyme disease patient addresses geocoded, Howard County, Maryland*

Year	Confirmed cases (n)	Probable cases (n)	Geocoded case addresses		
			Confirmed n (%)	Probable n (%)	Total n (%)
2001	68	-	62 (91)	-	62
2002	39	-	37 (95)	-	37
2003	42	-	39 (93)	-	39
2004	94	-	86 (91)	-	86
2005	150	-	137 (91)	-	137
2006	113	-	99 (88)	-	99
2007	358	-	326 (91)	-	326
2008	307	63	275 (90)	59 (94)	334
2009	203	65	184 (91)	62 (95)	246
2010	172	65	153 (89)	56 (86)	209
2011	137	58	123 (90)	55 (95)	178
Total	**1,683**	**251**	**1,521 (90)**	**232 (92)**	**1,753 (91)**

*Cases provided directly from the Maryland Department of Health and Mental Hygiene

Table 2.3. High-risk spatial clusters detected using two km radius maximum cluster size

Cluster	**Radius** (km)	**Observed case households** (n)	**Expected case households** (n)	**Relative risk**	**p-value**
1	1.67	109	46.0	2.46	<0.001
2	1.76	40	11.8	3.46	<0.001
3	1.97	42	14.4	2.98	0.001
4	1.76	80	38.9	2.11	0.002
5	1.58	20	4.8	4.24	0.030
6	0.15	8	0.6	12.94	0.030
7	0.31	13	2.0	6.41	0.040

Table 2.4. Percent of land area* according to land use and land cover classification, inside and outside of two km high-risk clusters, Howard County, Maryland

Variable		Percent area inside and outside of high-risk clusters			
		Inside	Outside	p-value** (one ha)	p-value** (one km^2)
Land use category	Very low-density residential, dominated by forest	7.7	3.7	<0.001	0.171
	Very low-density residential, dominated by agriculture	3.1	4.7	<0.001	0.610
	Low-density residential	30.3	16.7	<0.001	0.017
	Medium-density residential	13.0	9.8	<0.001	0.479
	High-density residential	1.6	3.1	0.136	0.889
	Urban/commercial	2.3	10.5	<0.001	0.067
	Herbaceous/agriculture	13.8	27.0	<0.001	0.040
	Forest	27.9	23.9	<0.001	0.528
Land cover category	Herbaceous/agriculture	35.0	51.1	<0.001	0.030
	Urban/residential	14.2	18.1	<0.001	0.492
	Red-white oak forest	38.9	15.3	<0.001	<0.001
	Chestnut oak forest	2.9	6.3	<0.001	0.334
	Other deciduous forest	4.2	6.0	<0.001	0.612
	Mixed forest	4.7	2.6	<0.001	0.401

***** Land covered by water was excluded

****** Chi-squared tests conducted using two different units of area as the basis of calculation (one hectare [ha] and one km^2)

Table 2.5. Area-weighted* socio-demographic factors inside and outside of two km maximum radius high-risk clusters

Variable	**Inside clusters**	**Outside clusters**
Median household income ($)	140,662	124,394
Per-capita income ($)	57,189	48,798
Median home value ($)	620,990	557,844
Population growth during 2000-2009 (%)	2.47	2.19
Population density (people per mi^2)	1,238	1,116
Percent of population with ≥ bachelor's degree	39.6%	35.6%
Percent of population < 15 years old	22.8%	22.5%
Median age (years)	41.7	40.5

*****2009 estimates per census block group

FIGURES

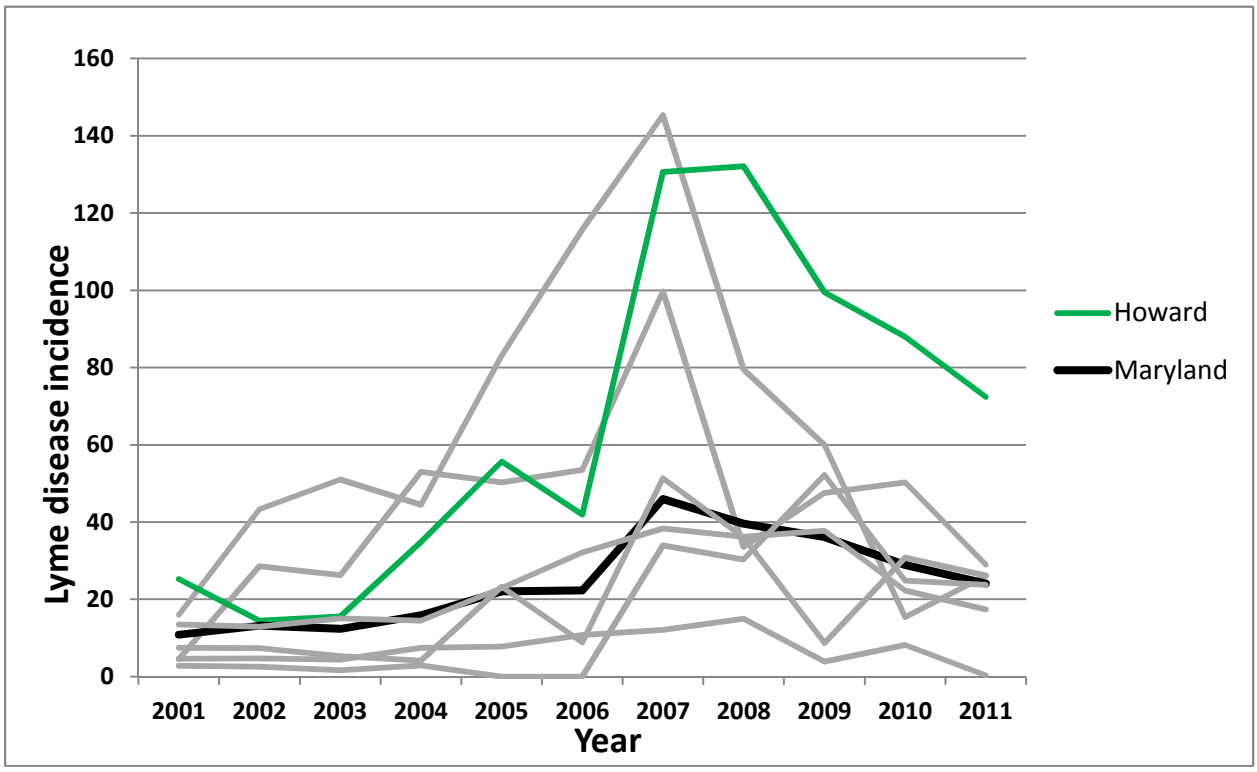


Figure 2.1. Lyme disease incidence in Howard County, neighboring counties, and Maryland, 2001-2011*

*Incidence was calculated as total (confirmed and probable) reported cases per 100,000 residents using 2005 U.S. Census Bureau estimates. Howard County incidence is displayed green, State of Maryland incidence is displayed black, and incidence in counties that share a border with Howard County are displayed in gray.

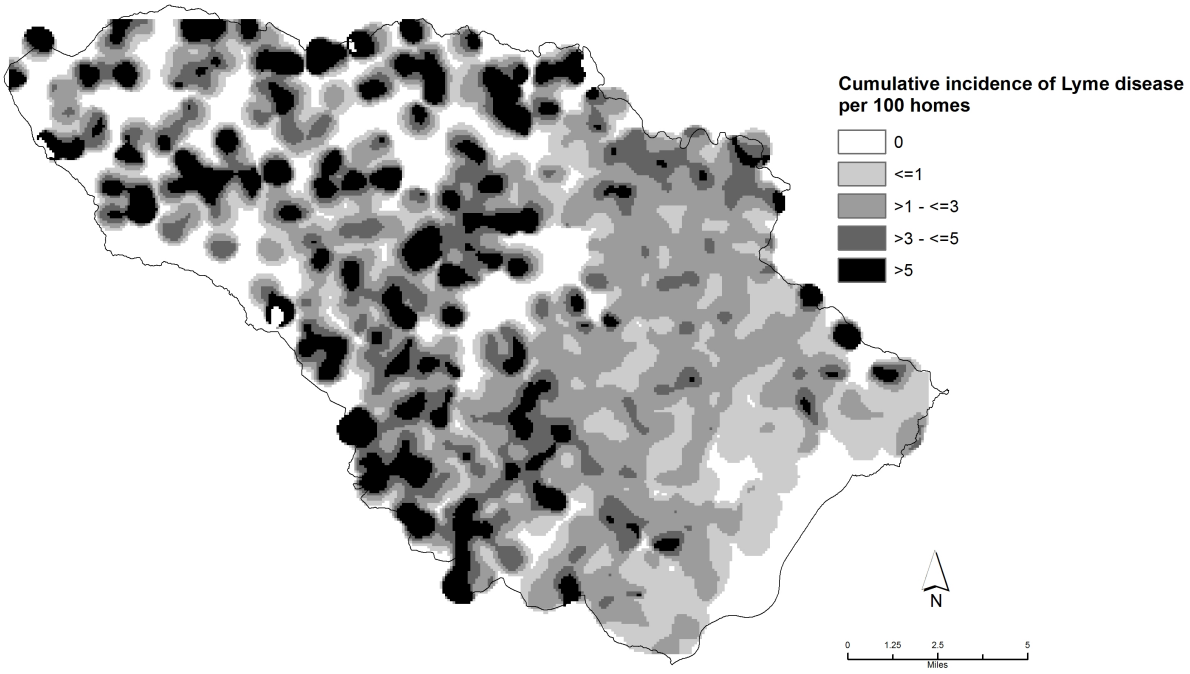


Figure 2.2. Smoothed case density surface of reported human Lyme disease in Howard County Maryland, 2001-2011*

*Map was created using kernel density functions with half mile bandwidth. Smoothed surface of case residences was divided by smoothed surface of all residences in the county to produce this map of cumulative incidence across space. Darker color indicates higher density of cases according to residence.

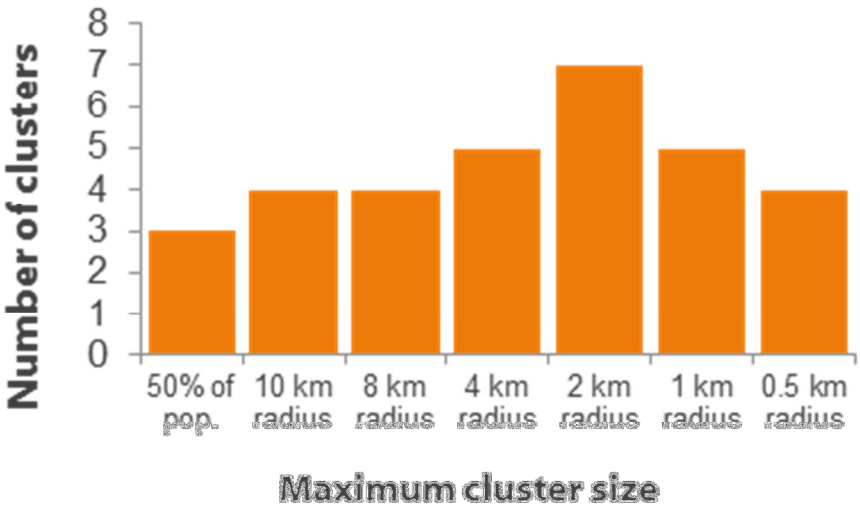


Figure 2.3. Number of high-risk clusters detected according to maximum cluster size limits

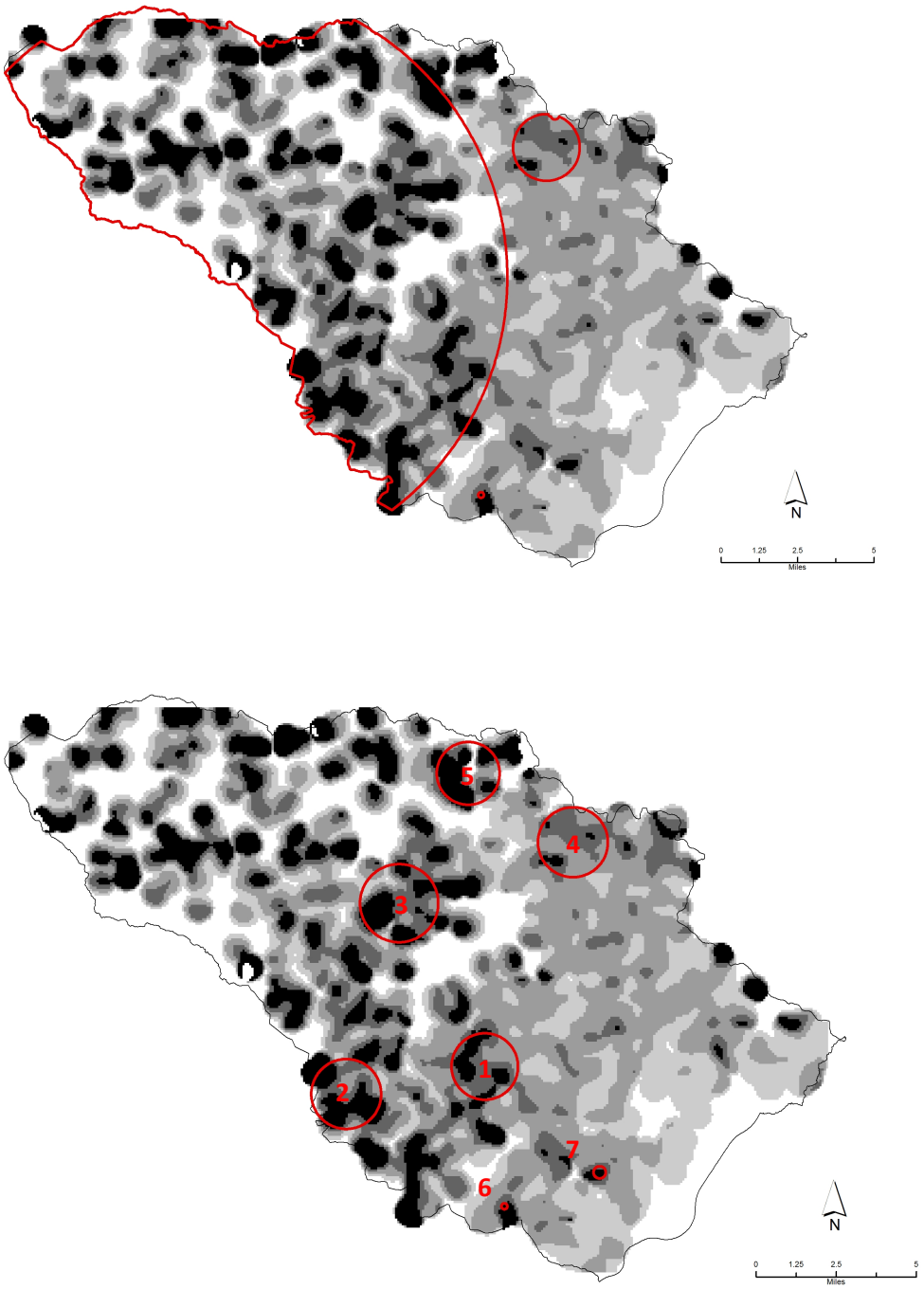


Figure 2.4. High-risk clusters detected given two software settings*

*Top panel displays the high-risk clusters as detected setting the maximum cluster size to the software default of 50% of the population. The bottom panel depicts those detected setting the maximum to two km. Clusters are numbered according to numbers in Table 2.3. High-risk clusters are depicted in red, and overlay the density surface displayed in Figure 2.2.

CHAPTER 3: PROJECT 2

3. HOUSEHOLD AND NEIGHBORHOOD CHARACTERISTICS AS RISK FACTORS FOR HUMAN LYME DISEASE

INTRODUCTION

Lyme disease is the most common vector-borne disease in the United States (Centers for Disease Control and Prevention 2013a), yet risk factors for human infection are not well-described. Substantial research effort has been directed at examining the spatial distribution as well as landscape and climatic factors associated with tick abundance and infection prevalence. Although these entomologic factors are necessary components for human Lyme disease risk, human infection depends not only upon abundance of infected ticks in the environment but also human behavior, movement, and interaction with the environment.

In the studies conducted to date that have examined risk factors for human Lyme disease, rural or low-density residential development has been positively associated with disease risk when compared to medium- and high-density development areas (Glass et al. 1995; Cromley et al. 1998; Orloski et al. 1998; Smith et al. 2001). The presence or close proximity of forests has been documented at various spatial scales as a primary risk factor for human Lyme disease (Maupin et al. 1991; Glass et al. 1995; Dister et al. 1997; Kitron and Kazmierczak 1997; Jackson et al. 2006a). Studies are mixed regarding a link between forest fragmentation and

human Lyme disease risk (Cromley et al. 1998; Allan et al. 2003; Brownstein et al. 2005; Jackson et al. 2006a; Killilea et al. 2008; Keesing et al. 2009; Diuk-Wasser et al. 2012).

While determination of individual-level risk factors for disease has been the mainstay of epidemiologic research, health scientists are increasingly aware of the importance of the context in which individuals reside on their behaviors and health outcomes (Diez-Roux 1998; Diez-Roux 2000; Diez Roux and Aiello 2005; Rothman et al. 2008). These "neighborhood" or "group" factors may augment or diminish one's specific individual risk of disease. Traditional epidemiologic study designs are either individual or group in nature (e.g., case-control and ecologic studies, respectively), and are unable to simultaneously examine the role of individual and group variables on an individual's risk of disease. Risk factors for Lyme disease may be individual behaviors or environmental factors that act at the household or neighborhood level. Additionally, neighborhood-level factors could affect both the abundance of infected ticks and the social context that influences how people move and interact with their environment.

Multilevel analysis (i.e., hierarchical, or mixed effect modeling) can not only account for the non-independence of observations within groups, but allows for description of both between- and within-group variability (Diez Roux and Aiello 2005). A multilevel epidemiologic analysis that can both quantify and examine the determinants of variability in disease risk across space has not been conducted for Lyme disease. In this multilevel analysis using data from a single county, environmental and socio-demographic associations with household-level risk of Lyme disease were examined. The goals of this project were to determine if Lyme disease risk was associated with any of the same environmental or socio-demographic factors

that were tied to high-risk clusters in Project 1, and to examine the relative contribution of household-level and census block group-level factors in describing spatial variation in risk.

METHODS

Study design

Associations between environmental and socio-demographic risk factors and household risk of human Lyme disease were assessed using an historical cohort of all households in one Lyme disease endemic county. Land use and land cover classifications were assessed at individual geocoded household points. Census block group of residence was considered the "neighborhood" for multilevel analysis. Ideally, a multilevel model for risk of Lyme disease should also include individual-level variables, although they were not ascertained as part of this study.

Study location and population

This study used data for all residences in Howard County, Maryland, which is located between Baltimore and Washington D.C. (Figure 3.1.). Howard County is 250 square miles and had a population in 2010 of approximately 287,000, with a density of 1,145 residents per square mile (US Census Bureau 2012). Households were enumerated using a geographic shapefile of all residential structures in Howard County as of 2004 (n=94,308), available from the Howard County Geographic Information Systems Division. Households present in Howard County in the 2004 geographic shapefile were assumed to be households present during the entire study period, 2001-2011.

Household-level Lyme disease classification

Potential Lyme disease cases that occurred during 2001-2011 were reported by clinicians and laboratories and subsequently investigated and classified according to standardized case definitions by local and state health officials. Cases occurring 2001-2007 were classified according to the case definition established in 1996 (Council of State and Territorial Epidemiologists 1996); cases occurring 2008-2011 were classified according a revised case definition that took effect in 2008 (Council of State and Territorial Epidemiologists 2008). For the purpose of this study, the Maryland Department of Health and Mental Hygiene (MDHMH) provided Lyme disease case information that was stripped of patient age and sex information, but contained residential address. Residential address data for reported patients were cleaned for misspellings and typographical errors with Google Maps and geocoded using the Howard County household address file within ArcGIS v.10.1 (Environmental Systems Research Institute, Redlands, CA).

The outcome assessed was reported Lyme disease in each household. All residences were classified according to a binary outcome of Lyme disease or no Lyme disease. Lyme disease classification was based on confirmed and probable cases reported to the health department during 2001-2011; households without a reported case of Lyme disease during the study period were considered to have no human Lyme disease. For all analyses, residence was considered to be location of infection. Use of patient information for this study was approved by Institutional Review Boards at the Centers for Disease Control and Prevention, MDHMH, and Colorado State University (Appendix 1.0.).

Explanatory variables and specifications

Potential explanatory variables for risk of Lyme disease were created from several datasets, and were chosen a priori, primarily based on previous research as described in Chapter 1. Land use and land cover category values were assigned to each house using ArcGIS according to the classification of the pixels in which each geocoded residence was located. Multilevel assignments (i.e., households nested within census block groups) were extracted according to year 2000 Census block group assignments using ArcGIS. Values for several variables were calculated or extracted according to census block group boundaries to classify the general area in which houses were located.

Household-level variables

Land use and land cover datasets were used to describe place of residence with regard to how general human use of the landscape and the dominant vegetation type. Two variables were ascertained at the household: land use classification and land cover classification; these variables were created using the same datasets as were used for Project 1.

For classification of land use, a dataset was obtained from the State of Maryland Department of Planning. This 2010 dataset was created from a combination of residential parcel data and aerial photography, with a stated resolution of one meter; the 24 land use categories in the Howard County dataset were collapsed into five classes for further analysis. Final land use categories were the following: low-density development (including very low-density development classes dominated by forest and agriculture, all lot sizes of at least one-half acre), medium-density development (lot sizes between one-eighth and one-half acre), high-

density or urban development (more than eight dwelling units per acre or other urban non-residential uses), herbaceous or agricultural use (e.g., parks, pastures, cropland, and brush), and forest (deciduous, evergreen or mixed forest types) (Table 3.1.).

For land cover, a dataset delineating dominant vegetation type was obtained from the United States Geological Survey Gap Analysis Program. This 2002 dataset was created from aerial video, satellite data, and ground assessments using the National Vegetation Classification System as the basis for the classification of dominant canopy species. Resolution was two hectares for land cover interpretation, but 30 meters for classified pixels. Land cover classes were initially collapsed into six classes as in Table 3.1, and then into four broad classes for regression analyses. Final land cover classes were grouped based on frequency and similarity with respect to potential role in the *B. burgdorferi* enzootic cycle and were the following: urban (impervious surfaces), herbaceous or agricultural (i.e., row crops, cultivated trees, pasture), red oak and white oak forest (oak-hickory forests with forest canopy dominated by *Quercus* [*alba, rubra*] and other oaks), and all other forest types (primarily other deciduous forests types as well rare deciduous/evergreen mixed forests).

Census block group-level variables

Forest fragmentation

Several indices of forest prevalence and fragmentation were calculated using Patch Analyst, the FRAGSTATS™ software add-in to ArcGIS (Rempel et al. 2012). Forest patch size and edge metrics were calculated using the forest class of the land use dataset described above; variables assessed were the following: proportion of census block group with land use

classification of forest; mean forest patch size in census block group; amount of forest edge relative to total census block group land area; amount of forest edge relative to amount of census block group classified as forest; and, mean amount of forest edge per forest patch in census block group (Table 3.1.).

Socio-demographic characteristics

In order to describe contextual effects related to income, age, and education, 2009 census block group estimates were assigned to all households based on their census block group (Table 3.1.).

Statistical analysis

Multiple statistical methods are routinely used to account for non-independence among observations, particularly mixed effect (or "multilevel") models and generalized estimating equations (Hanley et al. 2003; Waller and Gotway 2004; Hubbard et al. 2010; Subramanian and O'Malley 2010). Choice of which method to use depends on several factors, including whether correlation among observations is of interest in itself or just a statistical issue that should be handled (Hubbard et al. 2010).

Multilevel logistic regression models were created to examine the associations between household-level land use and land cover, census block group-level forest fragmentation and census block group-level socio-demographic factors and risk of Lyme disease. The multilevel analytic strategy was employed to account for the hierarchical nature of the independent variables and the potential non-independence of risk among households in the same census block group; moreover, this strategy was used to distinguish the relative importance of

household and census block group effects in explaining variation in risk of Lyme disease across census block groups.

Multilevel models were constructed with random intercepts for census block groups using PROC GLIMMIX with Laplace approximation in SAS v.9.3 software (SAS Institute, Cary, NC). Traditional intraclass correlation coefficients (ICC) are possible only in linear models; a pseudo-ICC was calculated using a "null" or "empty" model without any covariates, but with a random intercept for census block group (referred to as Model 0). The pseudo-ICC was used to describe the variation in risk of Lyme disease according to census block group. Specifically, the pseudo-ICC was calculated by dividing the estimate for the covariance parameter by the sum of the covariance parameter and the constant value of 3.29 (approximately $\pi^2/3$) (Snijders and Bosker 1999; Alexandrescu et al. 2011).

Univariate fixed effect logistic regression analyses were conducted to calculate crude odds ratios (OR) and for comparison to multilevel model findings. Three multivariable logistic regression models were created: 1) a model with only household-level characteristics (i.e., land use and land cover); 2) a model with only census block group-level variables (i.e., forest fragmentation and socio-demographic factors); and 3) a "full" model with both household and census block group effects. All multilevel models were additionally fit in models without random intercepts (single level models) to compare fixed effect estimates. Crude and adjusted odds ratios (aOR), 95% confidence intervals (CI) and p-values are presented for all models. Akaike's Information Criterion (AIC) values were used to assess goodness-of-fit (Akaike 1974).

Pseudo-ICC and covariance parameter values from each multilevel model were compared to examine how household and census block group effects reduced correlation within census block groups (i.e., between-block group variation), and determine how much correlation remained within census block groups after accounting for these covariates. The reduction in between-block group variance across models was assessed by calculating the proportion of the covariance parameter reduced given a set of fixed effects (Snijders and Bosker 1999; Alexandrescu et al. 2011; Bertille et al. 2013; Singh et al. 2013).

Household land use and land cover were entered in all models as categorical variables with urban land use and land cover considered the reference groups; census block group forest fragmentation and socio-demographic indices were entered as continuous variables. Odds ratios for all continuous variables were calculated per interquartile range (IQR) increase. Correlation of variables as evidenced by impact on standard errors was considered when constructing multivariable models. Multiplicative interaction was assessed between land use class and land cover class at the household location using a cross-product term in logistic regression using a p-value cutoff of 0.2.

RESULTS

Incidence of Lyme disease in the study population

Of 1,934 confirmed and probable Lyme disease cases reported among Howard County residents during 2001-2011, 112 cases (6%) were excluded from further geographic analyses because they either lacked an address or had a post office box listed, and 52 (3%) were excluded because listed addresses were not located in Howard County. In total, 1,753 (91%) cases (1,521 confirmed, 232 probable) from 1,672 unique residences were successfully

geocoded; these 1,672 houses were classified as having the outcome. Of 92,636 geocoded houses considered to have no human Lyme disease during the study time frame, five were unable to be assigned to a Howard County census block group. Thus, the final dataset included 92,631 non-case households. All households were nested within 118 census block groups in Howard County.

Characteristics of study households

The characteristics of study households according to several explanatory factors are shown in Table 3.1.; univariate logistic fixed effect associations are shown in Table 3.2. There were several significant differences in univariate analyses. When compared to households without Lyme disease during the study period, households with Lyme disease occurred more often in low- or medium-density residential development or in forest, and with dominant vegetation classified as red and white oak forest. Homes with reported Lyme disease occurred more frequently in census block groups with proportionally more area classified as forest and larger contiguous forest patches (i.e., more crude forest edge habitat (more forest edge per forest patch), and lower forest edge per forest area ratio; Table 3.2.). In univariate models, higher Lyme disease risk was also associated with residence in census block groups characterized by lower population density, higher average income, higher average home value, older average age of residents, and higher average education.

Census block group variation

The -2 Log Likelihood values of a null (i.e., no covariates) fixed effect model and a null model with only a random intercept for census block group (Model 0) were compared by a

likelihood ratio test using a mixture of chi-squared distributions; this comparison revealed substantial variation occurred across census block groups in risk of Lyme disease during 2001-2011 ($p<0.0001$).

Multilevel model development

Model 1

In multilevel Model 1, which included only household land use and household land cover classifications, low-density residential development (aOR: 2.12; 95% CI: 1.76-2.57), medium-density residential development (aOR: 1.96; 95% CI: 1.68-2.30), and land use classified as forest (aOR: 1.40; 95% CI: 1.04-1.87) were associated with increased risk of Lyme disease when compared to high-density residential or urban development (Table 3.3.). Land use classified as herbaceous or agriculture was associated with lower disease risk (aOR: 0.67; 95% CI: 0.46-0.97) than high-density residential or urban areas. Only one residential land cover classification was associated with elevated human Lyme disease risk compared to urban areas with no vegetative cover: red and white oak forest (aOR: 1.45; 95% CI: 1.22-1.71; Table 3.3.).

Model 2

In multilevel Model 2, which included forest and socio-demographic indices assessed at the census block group level, increasing percent of the census block group classified as forest was associated with elevated risk of disease (aOR: 1.43; 95% CI: 1.12-1.84). Although significantly associated with disease risk in univariate models, no other forest fragmentation metrics remained associated with Lyme disease risk in the multivariable setting, nor was there a clear trend in effect estimates (Table 3.4.). Mean amount of forest edge per forest patch was

not included in Model 2 due to correlation with other forest fragmentation variables and less biological relevance than the other forest variables. Although other forest variables were related to each other, inclusion of all variables did not substantially affect model stability and thus all were retained.

Income variables (median household income, median per capita income and median home value) were restricted to one variable due to multicollinearity. Median home value was selected not only because of its strong univariate association, but because in contrast to income itself, home value could better represent the exterior neighborhood landscape. After adjusting for census block group forest variables, residing within block groups with lower population growth (IQR aOR: 0.81; 95% CI: 0.73-0.91) and higher average home value (IQR aOR: 1.56; 95% CI: 1.31-1.86) were associated with elevated Lyme disease risk (Table 3.4.). In contrast, population density was not associated with household disease risk after controlling for the effect of other census block group socio-demographic and forest variables on household disease risk in a multilevel model (Table 3.4.).

Model 3

Model 3 combined all covariates in Models 1 and 2 (Table 3.5.). After accounting for forest and socio-demographic variables assessed at the census block group level, the same household-level land use and land cover characteristics remained significantly associated with elevated risk of Lyme disease in multilevel analysis. Low-density residential development (aOR: 1.85; 95% CI: 1.52-2.26) and medium-density residential development (aOR: 1.80; 95% CI: 1.54-2.12) were associated with increased risk of disease compared to high-density residential and

urban development (Table 3.5.). The positive association between residing in a home with land use classified as forest and disease was attenuated after adjusting for census block group forest prevalence and edge characteristics in this multilevel model (aOR: 1.29; 95% CI: 0.96-1.77; Table 3.5.). Residing in a location with land use classified as herbaceous or agricultural remained protective when compared to high-density residential or urban development (aOR: 0.62; 95% CI: 0.43-0.90). Likewise, the association remained between elevated risk of Lyme disease and residence in an area classified as red and white oak forest (compared to urban or minimal vegetative cover) after controlling for census block group forest variables and socio-demographic differences (aOR: 1.32; 95% CI: 1.11-1.57; Table 3.5.). Residence in homes located within pixels classified as red and white oak forest was also associated with elevated risk when specifically compared to residence in all other forest classes (aOR: 1.35; 95% CI: 1.1-1.7).

In multilevel Model 3, after adjusting for differences between land use and land cover of households, increasing proportion of the census block group classified as forest remained positively associated with disease (Table 3.5.) while other forest fragmentation variables were not significantly associated with Lyme disease risk (Table 3.5.). Similarly, positive associations between disease risk and lower population growth and higher average census block group home value remained consistent after adjustment for household-level characteristics. There was no evidence of interaction on a multiplicative scale between household land use class and land cover class with respect to risk of Lyme disease (p=0.98). Household land use and land cover were not subject to substantial confounding by variables assessed at the census block group level. Likewise, effect estimates of the census block group variables changed minimally after controlling for household land use and land cover classifications.

Relative contribution of levels in explaining spatial variation in risk of Lyme disease

The pseudo-ICC of Model 0, the null model without covariates, indicated that 8.3% of variation in risk of Lyme disease in Howard County was due to census block group of residence (Table 3.6.). This same value can also be conceived of as the degree of correlation, or "non-independence", in risk among homes within each census block group. The variance between block groups decreased after taking into account the covariates in Model 1; specifically, 39.7% of the total variance between block groups was explained by household specific land use and land cover characteristics. The census block group-level fixed effects included in Model 2 explained a much higher proportion of variance (63.9%) in risk of Lyme disease between block groups. When these variables were together in Model 3, the most variation across census block groups was explained; however, the results were not additive (Table 3.6.). Specifically, household land use and land cover, census block group forest characterization, and census block group socio-demographics explained 73.5% of the variance in risk between census block groups, leaving over 25% of the variance in risk between census block group still unexplained. Model 3 displayed the lowest AIC value and was considered the best of the candidate models (Table 3.6.).

In Figure 3.2., census block groups are displayed according to whether there was higher than expected variance, lower than expected, or similar to expected (at α=0.05), given each model described above. As covariates were added to a null model, much of the variation between block groups disappeared. Four block groups displayed higher than expected variation in risk and four block groups displayed lower than expected variation in risk given the Model 3 covariates. Thus, the covariates explained most of the spatial variation in risk of Lyme disease

according to census block group in Howard County, yet a substantial portion was unexplained, particularly in these areas, and may be accounted for by unmeasured household characteristics or individual behavioral differences.

Sensitivity analyses

To assess the relative contribution of the two groups of census block group variables (forest characteristics and socio-demographic variables), Models 2 and 3 were replicated using each group of variables separately (Appendix 2.1.). The socio-demographic variables explained more of the variation between census block groups than did forest variables; however, the model with forest variables only displayed better model fit. Nevertheless, both groups of variables in models together explained the most variation across census block groups and fit the data best, as evidenced by a lower AIC value.

Multivariable model 3 was also run in iterations of only two forest metrics at a time—proportion of the census block group classified as forest and one other. This analysis was conducted to ensure that possible associations were not attenuated by over-adjusting for related forest variables. Model fixed effects were essentially unchanged.

Census block group forest indices were also calculated using red and white oak forest. Univariate and multivariable findings (in Models 2 and 3) were similar in direction to those calculated using the land use forest class, but no metrics were significantly associated with disease (Appendix 2.2.). Risk of disease remained elevated for household residence in red and white oak forest when compared to residence in an urban area (without substantial vegetation) after adjusting for the structure and prevalence of red and white oak forests in the census block

group. The AIC value of this Model 3 was the lowest of any, indicating the best model fit; however, less variation in risk across space was explained by using red and white oak to calculate forest fragmentation compared to land use classified as forest.

The robustness of study findings were assessed with regard to choices made in grouping land use and land cover classes, and to missing data. Overall, these choices had minimal effect on findings; nevertheless, upon reclassification, residence in land use classified as forest became more highly associated with disease, and the protective effect of residence in land use classified as herbaceous or agricultural diminished (Appendices 2.3.-2.5.).

Fixed effect estimates of multilevel Model 3 were compared to the same model implemented as a generalized estimating equations (GEE) model with an exchangeable correlation structure. The final associations in Model 3 were mostly robust to the choice of modeling method, except that associations with population growth and percent of the population with a bachelor's degree were diminished in the GEE model (Appendix 2.6.).

Comparison of associations in multilevel and single level models

If non-independence between observations in this study was not taken into account, inference on associations with human Lyme disease risk would have been based on results from single-level models in Tables 3.3-3.5. In this analysis, although effect estimates were slightly diminished in magnitude and confidence intervals were wider after accounting for correlation in the multilevel model structure, strong associations with disease risk would have been detectable regardless of method.

DISCUSSION

Risk of Lyme disease in endemic areas may be tied to factors at individual, household, and neighborhood levels (Glass et al. 1995; Orloski et al. 1998; Smith et al. 2001; Connally et al. 2009). This project was the first known multilevel analysis of risk factors for human Lyme disease and the first study that examined the environmental and socio-demographic associations with reported Lyme disease among all households in one endemic county. Household risk of Lyme disease was associated with residence in suburban forested environments of low- and medium- density development characterized by red and white oak forest cover and in areas with more contiguous forest. In this analysis, some variation in human disease risk was due to census block group of residence, but most variation in risk occurred within census block groups. This finding underscores the fine-scale spatial variation in human disease risk. Use of multilevel model structure allowed 1) accurate simultaneous assessment of the associations between household and census block group environmental and socio-demographic characteristics with household risk of Lyme disease; 2) quantification of the variation in risk between census block groups; and 3) assessment of the relative contribution of household and census block group characteristics to reduction in the between census block group variation in risk of Lyme disease.

Despite use of different methodology from that in Project 1, this analysis also demonstrated variation in risk of Lyme disease across space. Specifically, 8% of variation in risk of Lyme disease across space was due to solely to census block group of residence. Most of the variation across census block groups could be accounted for with household land use and land cover classification and census block group-level differences in forest and socio-demographics,

but some unexplained variation across space remained. Most variation (92%) in risk of Lyme disease between houses was not due to geographic "neighborhood" location, but to other factors that vary within neighborhoods. Such fine-scale variation in risk could be a product of household landscape characteristics, individual behaviors, and chance. Ignoring the nested structure of households within census block groups could have resulted in erroneously small standard errors for coefficient estimates (Hanley et al. 2003; Waller and Gotway 2004; Diez Roux and Aiello 2005); however, this would have had relatively minimal impact on broad conclusions in this study. Residual between-census block group correlation remained after accounting for available covariates. Further examination into the effect of this correlation on fixed effect associations interpreted on a different level (household) and need for spatially-structured random effects is necessary.

Risk of reported Lyme disease was associated with low- and medium-density residential development, residence in areas with dominant vegetation classified as red and white oak forest, proportion of the census block group classified as contiguous forest, and residence in wealthier and more educated census block groups with lower population growth during the study period. These findings, which were obtained with a novel analysis method and study population, support many findings already documented in the scientific literature. Risk of Lyme disease has been associated with residence in lower-density development in studies conducted in Maryland (Glass et al. 1995), New Jersey (Orloski et al. 1998), Pennsylvania (Smith et al. 2001) and Connecticut (Cromley et al. 1998). Cromley et al. (1998) found rural areas to be at higher risk than medium-density development areas in a 12 town region of Connecticut in a geographic information system (GIS) analysis. One other published GIS-based study utilized

residential-level Lyme disease patient information within a single county to analyze environmental associations with Lyme disease risk. That study was conducted in Baltimore County, Maryland, which borders Howard County, and the authors found "highly developed" land to be associated with lower risk than all other development classes (Glass et al. 1995). Self-reported "rural" residence was associated with 14 times higher odds of Lyme disease (95% CI: 1.7-116.4) as compared to self-reported suburban or urban residence classification in Hunterdon County, New Jersey (Orloski et al. 1998). In addition, self-reported rural residence was associated with three times higher odds of Lyme disease than urban residence in Chester County, Pennsylvania (Smith et al. 2001). The mechanism through which residing in lower-density residential development impacts human risk is not clear. This association could be a surrogate for living within closer proximity to forest, more residential properties present that contain or border forests, or more forest-associated outdoor recreation than other areas.

In Howard County, residence in forest dominated by red and white oak species was associated with elevated risk of disease in Howard County. Although blacklegged tick abundance has been tied to oak and maple forests, Projects 1 and 2 are the first known studies to associate a specific forest type with human disease risk in the eastern United States. The importance of red and white oak forests as compared to other oak forests is unknown, but could be linked to the corresponding forest understory, abundance of leaf litter, underlying soil characteristics, or increased acorn production. Acorns produced by oak trees are an important source of food for small rodents, including white-footed mice. Abundance of reservoir competent white-footed mice and other small mammals are linked to availability of nutritious food that promotes winter survival and reproduction (Jones et al. 1998a; Ostfeld 2011). Acorn

masts have been linked to abundance of reservoir hosts, ticks, and elevated *Borrelia burgdorferi* infection prevalence (Jones et al. 1998b; Ostfeld et al. 2001; Ostfeld et al. 2006). Nevertheless, the association with residence in an area dominated by red and white oak forest may solely reflect residence in deciduous forest in general, as red and white oak forest is the most common type of deciduous forest in Howard County. Future effort should be directed at elucidating the potential association of red and white oak forests with elevated risk. Examination of soil as an explanatory variable will help tease apart the importance of soil underlying the forest from the forest composition itself. GIS datasets have become more refined in recent years, as has computing power, which will promote specificity in understanding of the mechanism through which land cover and forest types are associated with disease risk.

Of all forest indices assessed at the census block group level, only the proportion of the census block group classified as forest remained significantly associated with human illness after adjusting for household environmental characteristics and census block group differences in socio-demographic characteristics. Presence of forest near homes has been associated with tick abundance and human risk, although the proportion of land in a census block group classified as forest has not previously been evaluated with respect to human risk (Maupin et al. 1991; Glass et al. 1995; Dister et al. 1997; Jackson et al. 2006a). This finding underscores the potential importance of the broader neighborhood ecosystem structure in determining human risk, independent of the environmental conditions at the home itself.

Associations with forest patch size and edge metrics differed between Project 1 and Project 2, despite use of the same landscape fragmentation calculation method. Specifically, high-risk clusters identified in Project 1 were characterized by proportionally more edge and smaller forest patches. In contrast, this project demonstrated that homes with reported human Lyme disease were more often in areas with larger forest patches and more total forest edge. Nevertheless, those associations disappeared in multivariable models with proportion of the census block group classified as forest, suggesting volume of forest may be more important than forest structure with respect to human disease risk. Association between larger forest patches and increased human disease risk was also found by Brownstein et al. (2005) and Diuk-Wasser et al. (2012), but Allan et al. (2003) found lower blacklegged tick abundance in larger forest patches. Jackson et al. (2006) identified a specific forest to herbaceous cover edge index to be associated with increased of human Lyme disease incidence in road bounded polygons (Jackson et al. 2006a); whereas, in the current analysis type of edge was unspecified. Conflicting results in the literature may result from varied fragmentation measurement methods, different spatial units of aggregation, different outcome measures (human vs. entomologic), and univariate vs. multivariable analysis methods. In this study, forest fragmentation was calculated using land use classified as forest, which inherently is land classified as not belonging to predominantly residential areas. This measurement method was unable to address more fine-scale forest structure within census block groups or among smaller patches. The role that landscape fragmentation plays in increasing human risk remains to be elucidated, although these fragmentation indices did explain some portion of the variation in risk across space in Howard County.

Household Lyme disease risk was higher in census block groups with lower population growth and higher average home value. Lower population growth may indicate a more stable forest ecosystem, rather than one subject to clearing for development. The impact of forest clearing on disease risk has not been examined; however, areas with active or recent construction may be avoided for recreational purposes as compared to pristine forest. Jackson and colleagues (2006) published the only other study to include median income in census block group. Their goal was solely to control for potential differences in access to health care, rather than to examine income as a potential contextual-derived variable associated with reported Lyme disease (Jackson et al. 2006a). Clearly, home value itself does not predict Lyme disease risk, but is a surrogate for other landscape and behavioral characteristics that are difficult to ascertain, possibly including whether residential land includes or is bordered by forest.

Land use and land cover are often lumped together in datasets frequently used in GIS-based analyses. Yet, these two mechanisms by which to classify the landscape may be independently relevant to Lyme disease risk. Specifically, analyzing land cover helps understand how dominant vegetation types may foster the enzootic cycle. In contrast, land use data provide information with respect to the dominant purpose of the land. This analysis demonstrated an absence of a synergistic or antagonistic multiplicative effect between land use and land cover types at the household and human disease risk. Nevertheless, synergism of these effects is still possible on an additive scale.

Group-level variables considered in multilevel analysis can be conceived of as integral variables or derived variables (Diez Roux 2002; Diez Roux and Aiello 2005). Integral variables

describe a characteristic of the group itself, whereas derived variables represent aggregation of individual-level data. In our analysis, forest indices assessed at the census block group are not only integral variables, but can be conceived of as biologic risk factors, as they may directly modify the risk in the environment. In contrast, the socio-demographic variables assessed at the census block group are integral variables, and though may be explanatory factors, are not biologic risk factors. They are surrogates for the context in which people reside, and may represent differences in behavior and interaction with the environment.

Limitations

This study was subject to several limitations. Misclassification of households with respect to the outcome of Lyme disease during the study period could have occurred through reporting and surveillance practices, as well as geocoding limitations described in Chapter 2. This potential misclassification is unlikely differential with regard to environmentally-derived exposure variables; however, outcome misclassification could be differential across space due to differences in clinician reporting practices. If people residing in specific parts of the county were more likely to seek health care from a clinician more inclined to report cases of Lyme disease to the health department, differential outcome misclassification across space could occur. Furthermore, differential outcome misclassification could be linked to higher income and education. If wealthier and better educated people had better access to medical care and were more likely to have a case of Lyme disease both diagnosed and reported, this could serve to overestimate the association of these variables with Lyme disease. Nevertheless, by controlling for the differences in average income and education across census block groups, we may have minimized the impact of this potential bias away from a null association.

Misclassification of explanatory environmental and socio-demographic variables was possible. In these analyses, location of residence was assumed to be the location of infection, and may not be accurate for some cases. Misclassification of exposure location was likely non-differential across space, resulting in a presumed bias toward null associations with disease risk according to location of residence. Moreover, significant associations identified in this project may have otherwise been stronger absent this type of misclassification, assuming common epidemiologic knowledge regarding directionality of bias applies to spatial analyses (Rothman et al. 2008).

Misclassification of land use and land cover assessed at the household could have occurred in several ways, including misclassification due to spatial error in geocoded house placement, consideration of the land use and land cover class at only one point location (pixel), and error in the creation of the datasets themselves. Land use pixel classifications were inherently determined in consideration of the broader context of the neighboring environment (i.e., low-density development as a class does not specifically represent density at only one pixel). Consequently, placement error of a house into a neighboring pixel would likely have had resulted in minimal bias. Presumably, this misclassification would be non-differential with respect to Lyme disease status; however, because there are more than two categories for land use and land cover, potential direction of non-differential bias is unknown (Rothman et al. 2008). Measurement error in census block group assignment and in calculation of forest indices were also possible, and would also likely be non-differential with respect to disease status.

Assumptions of multilevel (or mixed effect, hierarchical or conditional) models can be difficult to verify (Hubbard et al. 2010). Population average (or "marginal", or GEE) approaches to dealing with non-independence between observations are generally more robust to misspecifications (Hanley et al. 2003; Hubbard et al. 2010). Nevertheless, effect estimates from the final multilevel model were compared to those from a GEE model; overall findings were not substantially different.

Multilevel analysis can be a powerful tool in epidemiology, but is frequently limited by lack of foresight to collect appropriate data at different levels and lack of understanding of the most appropriate groups to model (Diez-Roux 2000). In this study, census block group data were not collected specifically to assess census block group effects in association with Lyme disease, and thus our ability to define census block group-level effects for Lyme disease was hampered. Census block group was used as the neighborhood in this study. Unfortunately, the neighborhood context in which people reside, how it affects Lyme disease risk, and variability across space is likely best described by other (unknown) boundaries. Lastly, although ideal multilevel models of Lyme disease risk in an endemic area would include three levels (individual, household, and neighborhood), this analysis lacked individual-level data. Individual differences in behavior (e.g., propensity for gardening, jogging in woods, or preventive measures used) or more specific residential landscape characteristics (e.g., property size or forest border) likely account for most variation in risk of Lyme disease but were absent from this analysis.

Strengths

This study was the first to assess environmental and socio-demographic factors with risk of human Lyme disease on a sub-county scale. The study population consisted of all households in one endemic county, and was thus not subject to substantial selection bias. This robust sample size also contributed to high power in a multilevel analysis. Under-powered studies can be a problem; power in multilevel analysis depends on the number of groups and the number of observations within each group. Simulation studies have demonstrated that a minimum of 50 groups and 30 observations per group are needed to assure model convergence (Maas and Hox 2005; Moineddin et al. 2007).

Restriction of the study population to a single county minimized the impact of surveillance differences across counties and states that could occur as detailed in the Background and Project 1. Additionally, availability of household-specific information for non-case households allowed for more fine-scale analyses than possible with case count totals within administrative boundaries. Case count aggregation is typical with notifiable diseases and corresponding analysis is purely ecologic.

Conclusions

Without a human vaccine, greater understanding of the underlying causes of spatial variation in risk is necessary to identify and implement appropriate methods of disease prevention. Prevention of Lyme disease can theoretically occur at individual, household, or community levels. In this novel approach to understanding the epidemiology of human Lyme disease, the utility of multilevel analyses was evident. With respect to Lyme disease, multilevel

analyses have the ability to disentangle the importance of various levels with respect to risk. This type of analysis may help better understand the level at which various factors act on disease risk, and in turn inform the appropriateness of prevention methods that act on different levels.

This analysis was the first to apply to multilevel model structure to Lyme disease, a method more often used in the social sciences than the health sciences. Importantly, much multilevel research does not address spatial autocorrelation and its impact on fixed effect findings or on interpretability of findings at higher levels. Few authors have bridged the disciplines that commonly employ either multilevel analysis or spatial analysis; many methodologic questions remain (Chaix et al. 2005a; Chaix et al. 2005b; Merlo et al. 2005). One major question is of the best measure to evaluate the importance of the group variable, the Pseudo ICC as evaluated here, or the median odds ratio, a measure which translates the variance to the more interpretable odds ratio scale (Merlo et al. 2006).

Further analyses of risk factors for Lyme disease should be directed at defining what scale of "neighborhood" is most important in describing disease risk. For example, one could assess characteristics in buffers of increasing sizes around case and control homes, and determine where associations with disease risk are maximized. Moreover, it would be informative to examine random slopes in addition to random intercepts across neighborhoods to determine whether fixed effect associations differ across space or are relatively uniform. The possibility remains that risk factors are non-uniform in distribution or effect across space, even

within an endemic county; this variability could provide insight into why consistent factors have not been well-defined in the decades since the discovery of Lyme disease.

Although the household and census block group-assessed environmental and socio-demographic factors utilized as explanatory variables in this study explained some of the spatial variation in Lyme disease risk, unmeasured factors also contributed to Lyme disease risk. Additional assessment of socio-demographic factors such as income, education, and landscape factors obtained at the individual or household level is ideal. Some important factors may include more specific landscape design at each house (e.g., manicured space, sunny vs. shady areas, gardens), property size, and percent forest in a smaller "neighborhood", human outdoor behavior, and the unmeasured spatial variation in both reservoir hosts and tick abundance. Relatively few studies have examined variables that assess human interaction with the environment—most focus on landscape characteristics and use of personal protective measures (Orloski et al. 1998; Smith et al. 2001; Connally et al. 2009; Finch et al. 2014). These studies have demonstrated that increased time spent outside, doing yard work (including performing brush clearing activities), using tools (electrical, gasoline, or hand-powered), and having children who participate in outdoor sports are associated with increased Lyme disease risk. Broad, well-powered studies have not been conducted to better understand how human behavior patterns contribute to increased risk; such information could greatly help refine prevention messages.

TABLES

Table 3.1. Howard County households according household- and census block group-level characteristics and reported Lyme disease during 2001-2011

Variable type	Variable	Houses with Lyme disease* N=1,672	Houses without Lyme disease** N=92,631
Household-level variables			
Land use n (%)	Low-density development	565 (33.8)	19,694 (21.3)
	Medium-density development	738 (44.1)	35,802 (38.7)
	Urban/high-density development	270 (16.2)	28,390 (30.7)
	Herbaceous/agriculture	36 (2.2)	4,906 (5.3)
	Forest	63 (3.8)	3,834 (4.1)
Land cover n (%)	Urban	625 (37.4)	41,009 (44.3)
	Herbaceous/agriculture	608 (36.4)	32,641 (35.3)
	Red-white oak forest	271 (16.2)	9,009 (9.7)
	Chestnut oak forest†	80 (4.8)	5,428 (5.8)
	Other deciduous forest†	49 (2.9)	2,656 (2.9)
	Mixed deciduous/evergreen forest†	39 (2.3)	1,860 (2.0)
Census block group-level variables			
Forest fragmentation in census block group Mean (IQR)‡	Proportion classified as forest	0.272 (0.184)	0.245 (0.162)
	Mean forest patch size (km^2)	0.204 (0.187)	0.176 (0.138)
	Mean forest edge per forest patch (m)	3,069 (1,687)	2,835 (1,594)
	Mean forest edge per total land area (ratio*1,000)	1.359 (0.562)	1.339 (0.566)
	Mean forest edge per forest area (ratio*1,000)	5.93 (3.52)	6.50 (3.70)
Socio-demographics in census block group (2009 estimates) Mean (IQR) ‡	Median age (years)	39.8 (8.2)	38.8 (8.1)
	Proportion with ≥ bachelor's degree	0.389 (0.076)	0.370 (0.098)
	Population density (# people per mi^2)	2,206 (2,607)	2,674 (2,761)
	Population growth (% change 2000-2009)	1.77 (2.30)	1.85 (2.40)
	Per capita income ($)	48,767 (15,353)	45,230 (15,440)
	Median household income ($)	121,803(37,244)	111,705 (54,550)
	Median home value ($)	500,567 (273,858)	440,545 (255,611)

*n=1,649 for forest size and edge variables in census block groups (missing data=no pixels were classified as forest)

**n ranges between 91,449 for forest size and edge variables, 92,472 for socio-demographic indices (data unavailable for four census block groups, 92,626 and 92,603 for land use and land cover, respectively (some households in pixels with no data).

†Collapsed to one category ("other forest") for regression analyses

‡IQR = interquartile range

Table 3.2. Univariate single level fixed effect associations between land use, land cover, forest indices and socio-demographic indices and household risk of Lyme disease, Howard County, Maryland

Variable type	Variable	Variable specification*	Odds ratio	95% CI**	p-value
Land use		Low-density development	3.02	2.61-3.49	<0.001
		Medium-density development	2.17	1.88-2.49	<0.001
		Forest	1.73	1.31-2.28	0.001
		Herbaceous/agriculture	0.77	0.55-1.10	0.146
		Urban/high-density development	*ref*		
Land cover		Herbaceous/agriculture	1.22	1.09-1.37	0.001
		Red-white oak forest	1.97	1.71-2.28	<0.001
		Other forest	1.11	0.93-1.32	0.240
		Urban	*ref*		
Forest fragmentation indices	Percent area classified as forest	IQR (16.2% increase in percent of total land in census block group)	1.30	1.22-1.38	<0.001
	Forest edge per total land area	IQR (0.57 increase in ratio of edge to area*1,000)	1.05	0.99-1.11	0.088
	Forest edge per forest patch	IQR (1,594 m)	1.13	1.08-1.17	<0.001
	Mean forest patch size	IQR (0.14 km^2)	1.16	1.12-1.21	<0.001
	Forest edge per forest area	IQR (3.74 increase in ratio of edge to area*1,000)	0.76	0.71-0.82	<0.001
Socio-demographic indices	Population density	IQR (2,761 people per mi^2)	0.74	0.69-0.79	<0.001
	Per capita income	IQR ($15,314 increase)	1.58	1.47-1.69	<0.001
	Percent population growth 2000-2009	IQR (2.4% increase in growth)	0.96	0.92-1.02	0.165
	Median household income	IQR ($54,550)	1.70	1.56-1.85	<0.001
	Median home value	IQR ($255,611)	1.64	1.54-1.76	<0.001
	Median age	IQR (8.1 years)	1.33	1.23-1.43	<0.001
	Percent of population ≥ bachelor's degree	IQR (9.84% increase in % population)	1.23	1.17-1.30	<0.001

*Land use and land cover were each class variables with reference group indicated (ref); continuous variable associations displayed per interquartile range (IQR) unit increase

**95% confidence interval

Table 3.3. Multivariable Model 1: multilevel and single level (fixed effects only) associations between land use and land cover classification at the household location and household risk of Lyme disease, Howard County, Maryland

Variable type	Variable specification	Multilevel model			Single level model		
		aOR*	95% CI**	p-value	aOR*	95% CI**	p-value
Land use	Low-density development	2.12	1.76-2.57	<0.001	2.78	2.39-3.24	<0.001
	Medium-density development	1.96	1.68-2.30	<0.001	2.10	1.83-2.42	<0.001
	Forest	1.40	1.04-1.87	0.025	1.55	1.17-2.05	0.002
	Herbaceous/agriculture	0.67	0.46-0.97	0.033	0.74	0.52-1.05	0.092
	High-density development/urban	*ref*			*ref*		
Land cover	Herbaceous/agriculture	0.99	0.86-1.15	0.909	1.10	0.98-1.23	0.126
	Red-white oak forest	1.45	1.22-1.71	<0.001	1.60	1.38-1.86	<0.001
	Other forest	1.02	0.85-1.23	0.805	0.99	0.83-1.18	0.908
	Urban	*ref*			*ref*		

*Adjusted odds ratio, reference group indicated (ref)
**95% confidence interval

Table 3.4. Multivariable Model 2: multilevel and single level (fixed effects only) associations between census block group variables and household risk of Lyme disease, Howard County, Maryland

Variable type	Variable	Variable specification* interquartile range	Multilevel model			Single level model		
			aOR**	95% CI†	p-value	aOR**	95% CI†	p-value
Forest fragmentation indices	Percent classified as forest	16.2%	1.43	1.12-1.84	0.005	1.44	1.24-1.67	<0.001
	Mean forest patch size	0.14 km^2	0.98	0.87-1.10	0.673	0.98	0.92-1.05	0.529
	Forest edge per forest area	3.74 (ratio*1,000)	1.12	0.94-1.34	0.188	1.12	1.00-1.26	0.045
	Forest edge per total area	0.57 (ratio*1,000)	0.93	0.81-1.07	0.320	0.96	0.87-1.05	0.327
Socio-demographic indices	Population density	2,761 people	0.96	0.81-1.14	0.608	0.95	0.84-1.08	0.423
	Percent population growth	2.4%	0.81	0.73-0.91	<0.001	0.82	0.76-0.88	<0.001
	Median home value	$255,600	1.56	1.31-1.86	<0.001	1.52	1.37-1.68	<0.001
	Median age	8.1 years	1.03	0.86-1.22	0.779	1.03	0.92-1.15	0.634
	Percent ≥ bachelor's degree	9.8%	1.10	0.99-1.22	0.073	1.10	1.03-1.18	0.007

*Associations displayed per interquartile range (IQR) increase

**Adjusted odds ratio

†95% confidence interval

Table 3.5. Multivariable Model 3: multilevel and single level (fixed effects only) associations of both household and census block group variables and household risk of Lyme disease, Howard County, Maryland

	Variable	Variable specification*	Multilevel model			Single level model		
			aOR**	95% CI†	p-value	aOR**	95% CI†	p-value
Land use	Low-density development		1.85	1.52-2.26	<0.001	2.05	1.70-2.47	<0.001
	Medium-density development		1.80	1.54-2.12	<0.001	1.78	1.53-2.07	<0.001
	Forest		1.29	0.96-1.77	0.089	1.31	0.98-1.74	0.072
	Herbaceous/agriculture		0.62	0.43-0.90	0.012	0.63	0.44-0.91	0.013
	High-density development/urban		*ref*			*ref*		
Land cover	Herbaceous/agriculture		0.91	0.78-1.06	0.235	0.94	0.82-1.08	0.411
	Red-white oak forest		1.32	1.11-1.57	0.002	1.35	1.14-1.59	<0.001
	Other forest		0.98	0.81-1.17	0.788	0.95	0.79-1.10	0.536
	Urban		*ref*			*ref*		
Forest fragmentation indices	Percent classified as forest	IQR (16.2%)	1.36	1.07-1.73	0.012	1.30	1.11-1.51	<0.001
	Mean forest patch size	IQR (0.14 km^2)	0.93	0.83-1.04	0.201	0.98	0.91-1.05	0.560
	Forest edge per forest area	IQR (3.74 ratio*1000)	1.03	0.87-1.21	0.772	1.03	0.91-1.15	0.172
	Forest edge per total area	IQR (0.57 ratio*1000)	0.95	0.82-1.09	0.426	0.97	0.88-1.07	0.540
Socio-demographic indices	Population density	IQR (2,761 people)	1.02	0.86-1.22	0.790	1.02	0.90-1.17	0.721
	Percent population growth	IQR (2.4%)	0.87	0.78-0.97	0.010	0.87	0.81-0.93	<0.001
	Median home value	IQR ($255,600)	1.46	1.21-1.75	<0.001	1.37	1.22-1.55	<0.001
	Median age	IQR (8.1 years)	0.97	0.82-1.15	0.707	0.98	0.87-1.10	0.689
	Percent ≥ bachelor's degree	IQR (9.8%)	1.13	1.01-1.25	0.027	1.10	1.02-1.18	0.012

*****Land use and land cover were class variables with reference groups indicated (ref); continuous variable associations displayed per interquartile range (IQR) increase

**Adjusted odds ratio

† 95% confidence interval

Table 3.6. Covariance and fit of multilevel models of Lyme disease risk, Howard County, Maryland

Model characteristic	Model 0: random intercept only	Model 1: household only	Model 2: census block group only	Model 3: household + census block group
Covariance parameter	0.2966	0.1741	0.1006	0.0867
P-value of covariance parameter*	<0.0001	<0.0001	<0.0001	<0.0001
Pseudo-intraclass correlation coefficient	0.083	0.050	0.030	0.022
AIC**	16,800	16,384	16,245	16,118

*P-value of likelihood ratio test that covariance parameter is equal to 0

**Akaike's information criterion

FIGURES

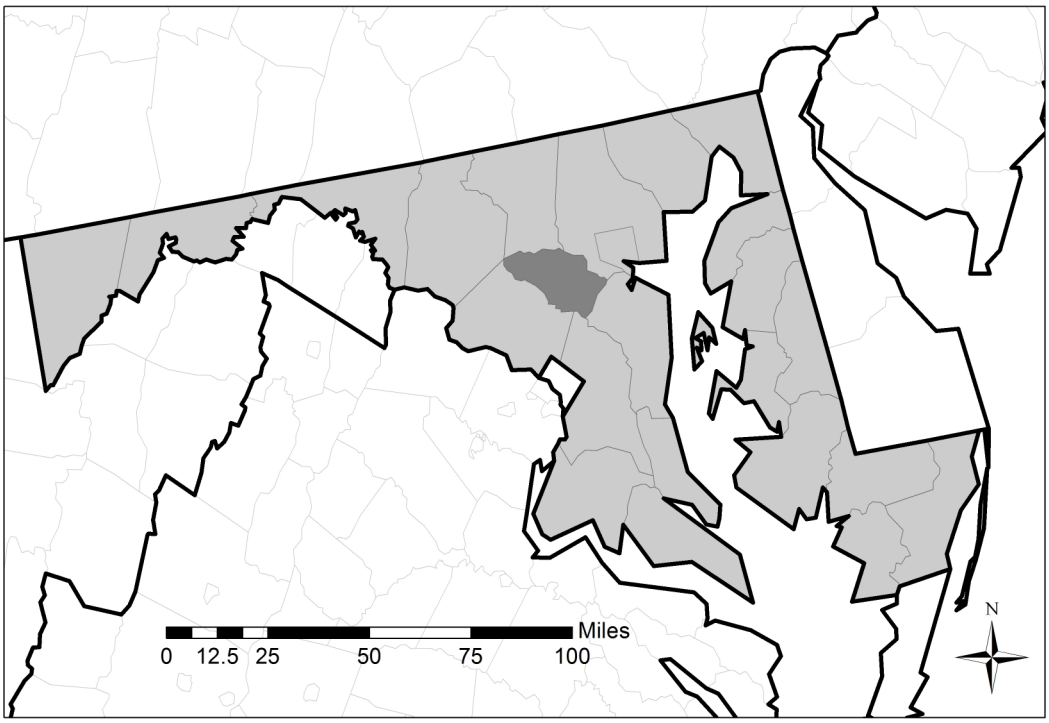


Figure 3.1. Map of Maryland with Howard County indicated by darker shading

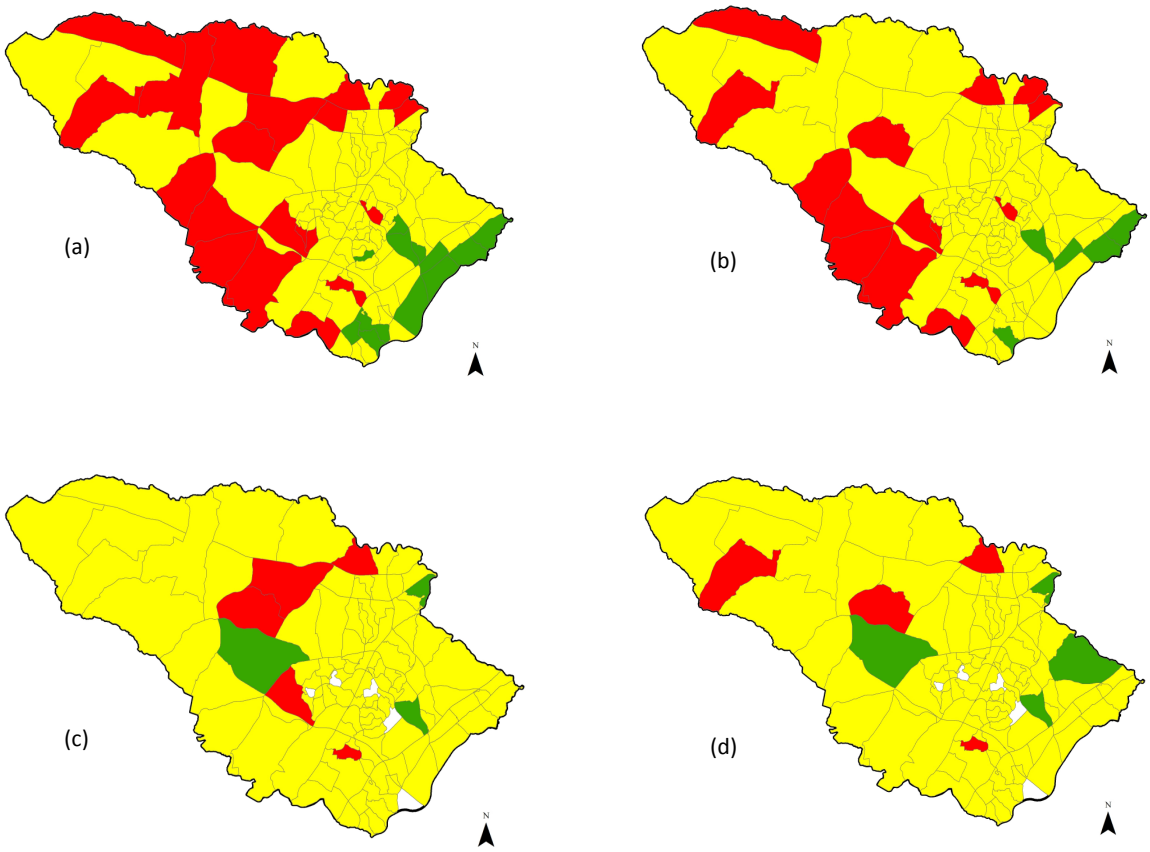


Figure 3.2. Residual variation in risk across census block groups given different covariates, Howard County, Maryland, 2001-2011*†

*Census block groups in red indicate significantly elevated variance ($p < 0.05$) of the random intercept compared to average, green significantly lower variance compared to average, and yellow, not statistically different than average. Panels depict this variation given a) null model with no covariates, b) Model 1 with household land use and land cover only, c) Model 2 with census block group covariates only, d) Model 3 with both household and block group covariates.

†Census block groups in white were excluded from analyses because 2009 estimates for socio-demographics were unavailable.

CHAPTER 4: PROJECT 3

4. DEER DENSITY, BLACKLEGGED TICK INFESTATION ON DEER, AND HUMAN LYME DISEASE RISK

INTRODUCTION

Over 30,000 cases of Lyme disease are reported each year in the United States; nearly all cases are from the northeastern, mid-Atlantic and north-central states (Centers for Disease Control and Prevention 2013a). Risk of acquiring Lyme disease varies across space even within endemic areas; a demonstration of this variation within one county is provided in Project 1. Reasons for fine-scale differences in the enzootic cycle or in human interaction with the cycle are not well understood (Wilson 1998; Killilea et al. 2008). An exploration of possible associations between land use, land cover, census block group-level landscape fragmentation, census block group socio-demographic factors and risk of Lyme disease in the same county was presented in Project 2.

In addition to the environmental factors examined in Projects 1 and 2, the abundance and infection prevalence of mammals and ticks in the environment are likely critical to variation in risk across space. White-tailed deer are an important host for adult *Ixodes scapularis* ticks and are critical to tick reproduction and geographic dispersal (Figure 4.1.) (Piesman et al. 1979; Wilson et al. 1990b; Barbour 1998). Accordingly, deer population reduction has been suggested as a possible Lyme disease prevention measure (Kilpatrick and Walter 1997; Hayes and Piesman

2003; Stafford 2007). Although deer are a preferred host for adult blacklegged ticks and most observational studies have shown some correlation between deer abundance and tick abundance, others have found no association (Piesman et al. 1979; Anderson and Magnarelli 1980; Schulze et al. 1984; Wilson et al. 1985; Wilson et al. 1990a; Daniels et al. 1993b; Stafford 1993; Duffy et al. 1994; Daniels and Fish 1995; Ginsberg and Zhioua 1999; Schulze et al. 2001a; Rand et al. 2003; Ginsberg et al. 2004; Jordan and Schulze 2005). Intervention studies of the impact of deer population reduction on tick abundance have demonstrated that near complete or complete elimination of deer in isolated settings will have a substantial impact on reproduction of blacklegged ticks (Wilson et al. 1984; Wilson et al. 1988; Rand et al. 2004). More moderate deer reduction efforts appear to have non-uniform effects on tick abundance (Deblinger et al. 1993; Stafford et al. 2003; Jordan et al. 2007). The effect of deer reduction may be buffered by increased tick abundance on remaining deer, but these complex data are difficult to generalize.

The associations between deer density and tick abundance with human Lyme disease risk are also not well-understood (Figure 4.1.). The role of deer density and tick abundance on deer with respect to human disease risk is separated by two years of blacklegged tick life cycle and multiple other factors, including density of infected reservoir hosts and density of nymphs, which contribute to ecological risk of disease. No studies have examined the spatial differences in deer density in association with human Lyme disease risk at any spatial scale. Currently, insufficient evidence exists to determine an association between deer density and human Lyme disease risk. Consequently, there is a lack of understanding as to whether moderate community deer management programs may have any effect on the spatial variation in risk of Lyme disease

in those areas. Abundance of blacklegged ticks on deer has been associated with human Lyme disease risk at the county level and between communities >20 km apart (Wallis et al. 1978; Kitron and Kazmierczak 1997). In areas with measured deer density and tick infestation on deer, evaluation of spatial associations between deer density, tick abundance, and human Lyme disease risk are possible, but have not been conducted on a sub-county scale.

As a complement to Projects 1 and 2, this project examines the role of deer density and tick infestation on deer in spatial variation in human Lyme disease risk in a single inland county. The goals of this project were to examine the associations between 1) deer density and tick infestation on deer, 2) deer density and risk of human Lyme disease, and 3) tick infestation on deer and risk of human Lyme disease.

METHODS

Study location and deer density estimates

Maryland consistently reports a high incidence of human Lyme disease (Bacon et al. 2008; Centers for Disease Control and Prevention 2013a). This study used data from Howard County, Maryland, which is located between Baltimore and Washington D.C. (Figure 4.2.). The approximately 250 square miles had a population in 2010 of approximately 287,000 (US Census Bureau 2012) and a substantial white-tailed deer population. The Howard County Deer Management Program has managed the deer population since 2000, although more informal management occurred prior to that time (Howard County Department of Recreation and Parks 2000). As part of the program, managed hunts and sharpshooting occurred each year at several public parks and natural areas to help curb white-tailed deer population growth within the County. To assess progress in reducing deer density, the management program annually

obtained deer density estimates during winter months in both hunted and un-hunted areas. These assessments were conducted using aerial Forward Looking Infrared Radar (FLIR) technology, and provided an annual snapshot of an inherently dynamic deer population. Deer density estimates in Howard County were first calculated in early 2001; estimates from parks across Howard County through early 2011 were examined for this analysis (Figure 4.2.).

Tick infestation on deer

Basic information for each deer culled as part of the deer management program (including age, sex, date and location of kill, and counts of ticks on the right ear) was extracted from paper records. Records pre-date the official beginning of the management program in 2000, as some deer removal occurred through the county government, in the early winter months of 1998, and Fall 1999-Winter 2000. Data on tick counts in years prior to deer density estimates were included to maximize sample size; analyses were conducted with cumulative rather than year-specific data. During the Fall 2012-Winter 2013 deer management season, in addition to counting the number of ticks on the right ear of each culled deer, ticks were placed in ethanol and shipped to the Centers for Disease Control and Prevention (CDC), Division of Vector-Borne Diseases in Fort Collins, Colorado, for species identification and pathogen testing. Pathogen testing was performed with a multiplex real-time polymerase chain reaction (PCR) assay designed to detect several tickborne pathogens (Hojgaard et al. 2014).

Household address and Lyme disease information

Potential Lyme disease cases that occurred during 2001-2011 were reported by clinicians and laboratories, and subsequently investigated and classified according to

standardized case definitions by local health officials. Cases occurring 2001-2007 were classified according to the case definition established in 1996 (Council of State and Territorial Epidemiologists 1996); cases occurring 2008-2011 were classified according a revised case definition that took effect in 2008 (Council of State and Territorial Epidemiologists 2008). For this study, the Maryland Department of Health and Mental Hygiene (MDHMH) provided Lyme disease case information that was stripped of patient age and sex, but contained residential address.

Address for all residential structures as of 2004 was available from the Howard County Geographic Information Systems Division (n=94,308). Households present in Howard County in the 2004 geographic shapefile were assumed to be the households present during the entire study period, 2001-2011. Residential addresses of patients were cleaned for misspellings and typographical errors with Google Maps and geocoded using the Howard County household address file within ArcGIS v.10.1 (Environmental Systems Research Institute, Redlands, CA). Household Lyme disease classification was based on confirmed and probable cases reported to the health department during 2001-2011; households without at least one reported case of Lyme disease during the study period were considered to have no human Lyme disease. Use of case information for this study was approved by Institutional Review Boards at the Centers for Disease Control and Prevention, MDHMH, and Colorado State University (Appendix 1.0).

Analytic methods

Descriptive analyses

Differences in the number of ticks found on the right ear of hunted deer were compared across parks and over time with Kruskal-Wallis tests, a non-parametric method for testing whether more than two samples arise from the same distribution. Difference in tick infestation by sex of deer was assessed by the Wilcoxon Rank-Sum test, a non-parametric method for testing whether two samples arise from the same distribution (Ott and Longnecker 2001).

Regression analyses

Regression models were used to examine associations between: 1) tick infestation on deer and deer density in Howard County parks; 2) deer density and human Lyme disease risk among homes surrounding parks; and 3) tick infestation on deer and human Lyme disease risk among homes surrounding parks.

Data restrictions

Regression models used information from areas of Howard County where relevant data were available. Deer density data were available only for specific parks and natural areas. Annual aerial estimates of deer density were captured on a single day and likely subject to substantial measurement error. To minimize the impact of this measurement error on analyses, analyses were restricted to parks with at least five years of deer density estimates during 2001-2011. Furthermore, a cumulative variable (i.e., median deer density during the study period) was used for analyses. A summary of several measurements may be a more meaningful

representation of deer density during the study period and should minimize the impact of a single measurement on findings.

To determine associations between tick infestation on deer and both deer density and human Lyme disease, parks were restricted to those with data regarding tick infestation on culled deer. Data were limited further to measurements on deer culled during October of any year. During October, questing behavior of adult blacklegged ticks peaks (Stafford 2007) and number of ticks on deer were most likely to be non-zero.

For analyses related to human Lyme disease risk, a half-mile buffer was created around the parks to capture all homes in close proximity to each park that may be "exposed" to the relative deer density and tick infestation in those parks. This buffer was chosen due to the relatively limited home range (≤ one mi^2) of white-tailed deer in suburban environments (Sparrowe and Springer 1970; Kilpatrick and Spohr 2000; New Jersey Department of Environmental Protection 2010).

Overview of regression models

For regression analysis to examine the association between tick infestation on hunted deer and deer density in each park, no specific direction of association was hypothesized because of conflicting information in the existing published literature. This analysis used individual hunted deer in Howard County as the unit of observation. Median deer density in a park was the independent variable and was examined both as a continuous and categorized variable. Tick counts on the right ear of individual deer were the outcome measure in Poisson and negative binomial regression models appropriate for count data.

For analysis of the association between deer density and human Lyme disease risk, the primary hypothesis was that human Lyme disease risk among homes surrounding parks would be higher around parks with higher average deer density. Individual households within the buffers around parks were used as the units of observation in logistic regression models. The primary deer density variable examined was categorized by quartiles of median density.

For models of tick infestation on deer and risk of human Lyme disease among homes surrounding parks, the primary hypothesis was increased odds of Lyme disease among homes with increasing tick infestation on deer in nearby parks. Mean tick infestation over time was the primary independent variable assessed; additionally, given lack of a priori knowledge of what measure of tick infestation may be most appropriate, other variables were examined. Logistic regression models were used to assess odds of Lyme disease among homes according to infestation of ticks on deer in the proximal park.

Selection of covariates

In the models used to evaluate the association between tick infestation on deer and deer density, the sex and the age of deer, and the year culled were included a priori. In human Lyme disease models, potential covariates included land use and land cover data, forest fragmentation indices, and census block group socio-demographic information as assessed in Projects 1 and 2. A covariate was considered a confounding variable if its presence changed the odds ratio of interest by 10% or more (Maldonado and Greenland 1993).

A dataset describing land use was obtained from the State of Maryland Department of Planning (resolution: one meter). A land cover dataset was obtained from the United States

Geological Survey Gap Analysis Program (resolution: 30 meters). Landscape fragmentation was calculated using Patch Analyst, a FRAGSTATS™ software add-in to ArcGIS (Rempel et al. 2012). Census block group data (including year 2000 block group boundaries, 2009 population estimates, 2009 estimates for income and home value variables, and 2000-2009 population growth estimates) were obtained from the Environmental Systems Research Institute (Redlands, CA).

Land use and land cover categories were collapsed into broad, biologically-relevant categories for use as potential covariates in analysis. Final land use categories were the following: low-density development (including very low-density development classes, in addition to areas of at least 90% single family or duplex dwellings, with lot sizes of at least one-half acre), medium-density development (lot sizes between one-eighth and one-half acre), high-density or urban development (more than eight dwelling units per acre or other urban non-residential uses), herbaceous or agricultural use (e.g., parks, pastures, cropland, and brush), and forest (deciduous, evergreen, or mixed forest types), herbaceous and agricultural (pasture, row crops, parks). Final land cover categories were the following: urban (impervious surfaces), herbaceous or agricultural (i.e., row crops, cultivated trees, pasture), red oak and white oak forest, and all other forest types (primarily other deciduous forests types as well rare deciduous/evergreen mixed forests). Environmental factors from land use and land cover datasets were extracted to ascertain classification at each geocoded household point location. Additionally, the Tabulate Area tool in ArcGIS was used to calculate the proportion of total area inside each park and buffer associated with each land use and land cover class.

Forest fragmentation indices were calculated in park and surrounding buffer areas; variables assessed were mean forest patch size and density of forest edge in relation to total area. Census block group variables including median home value, median household income, per capita income, population density, and population growth were assigned to each household based on its geolocated census block group assignment. Use of four-poster tick control devices (in place in a single park during most of the study period) was also examined as a potential covariate.

Model selection

For analysis of tick infestation on deer and deer density, models assessed were unadjusted, using different link functions appropriate for counts (i.e., Poisson, negative binomial), and adjusted for a priori covariates. For logistic regression models of human Lyme disease risk, final adjusted models were selected manually by assessing impact of potential confounders on the association of interest. Collinearity and instability caused by small cells were assessed and taken into consideration when selecting final adjusted models. Deviance residuals were examined to assess fit and identify outlying observations; logistic model fit was additionally assessed by Deviance, Pearson and Hosmer and Lemeshow goodness-of-fit tests (Hosmer and Lemeshow 2000; Hoffman 2004).

Due to the potential for non-independence of observations among deer hunted from the same park and human Lyme disease among households around each given park, the need to account for the potential correlation among observations was evaluated. Generalized estimating equations (GEE) with an exchangeable correlation structure were used for this

purpose; this method contrasts with the multilevel models presented in Project 2. Here, description of between-park variance was not of interest itself, only the fixed effect estimates. Thus, GEE was selected as an appropriate method to account for non-independence among observations (Hanley et al. 2003; Hubbard et al. 2010).

All models were fit using SAS software v.9.3 (SAS Institute, Cary, NC). Poisson and negative binomial models were implemented in the GENMOD procedure. Logistic models were constructed in the LOGISTIC procedure. All GEE models were implemented in the GENMOD procedure with a repeated statement, using an exchangeable correlation structure between observations associated with each park. Quasi-likelihood criterion (QIC) (Pan 2001) and covariance parameters were compared between GEE models fit with exchangeable correlation structures and those with independent structures to assess need for GEE models. Crude and adjusted odds ratios (OR and aOR, respectively), 95% confidence intervals (CI), and p-values were calculated.

RESULTS

Descriptive analyses

Deer density

A total of 15 parks (or natural areas) had aerial deer density estimates for at least five years during 2001-2011 (Tables 4.1. and 4.2.). Total deer density in all measured areas of Howard County ranged from a high of 94.3 deer per mi^2 to a low of 9.9 deer per mi^2. No obvious increase or decrease in density over time was evident (Table 4.1.). Area assessed at each site ranged from 0.17 mi^2 to 5.35 mi^2. Density estimates were highly variable not only

between parks but also across years within individual parks (Table 4.2.). The low total county density estimate in 2011 was likely due to the lack of an aerial estimate from one park with very overabundant deer that had been included in years past, rather than a true county-wide decrease in deer density (Table 4.1.).

Deer culled through deer management program

From early Winter 1998 through the Fall 2012 - Winter 2013 (except for Fall 1998 - Winter 1999 when deer management did not occur), 2,603 deer were culled in Howard County (Tables 4.3. through 4.6.). The number of deer removed ranged from 50 in the abbreviated first season to 256 in Fall 2000 - Winter 2001 (Table 4.3.). Nearly two-thirds (1,642/2,603; 63%) of all removed deer were female (annual range: 58% - 76% female). Of 2,588 culled deer with approximate age information, the mean age was approximately two years (median age: one and a half years). During this time 11 different public natural areas or parks were hunted; seven were areas with routine deer density measurement and accounted for 2,487 (96%) of all deer culled during the study period (Tables 4.3. and 4.4.).

Tick infestation on culled deer

Over 15 deer management seasons, 2,353 ticks were counted on the right ear of culled deer (Table 4.3.). Nearly all ticks (n=2,247, 95.5%) were counted on deer culled in parks with deer density estimates. Cumulatively, there were 0.90 ticks per right ear of culled deer (mean: 0.92; median: 0; range: 0 - 37 ticks per deer; variance: 7.42). Ticks per culled deer per park ranged from 0.32 - 1.07 (Table 4.4.). Tick infestation on deer differed across parks ($p<0.001$). When analyses were restricted to only those deer hunted during October, the peak of the adult

blacklegged tick questing season, the means across parks ranged from a low of 0.73 to a high of 2.63 ticks per hunted deer, but the difference was not statistically significant (p=0.152; Table 4.6.). Tick infestation on deer also differed across hunting seasons (p=0.013). During October, the mean number of ticks per deer varied between years from a low of 0.33 in 2005 to a high of 3.38 in 2011, with no apparent trend increasing or decreasing trend over time (Table 4.5.). The variance in tick abundance by park during October ranged from 2.1 to 41.5 (Table 4.6.); maximum ticks per deer by park, during October, ranged from seven to 37. Ticks were two times more abundant on male deer as compared to female deer (mean: 1.3 per male deer vs. 0.7 per female deer; p<0.001).

Borrelia burgdorferi *infection prevalence in ticks -- 2012-2013 season*

All but one tick collected from the right ear of hunted deer during the Fall 2012-Winter 2013 management season were adult *Ixodes scapularis;* one adult *Amblyomma americanum* was identified late in the season (February). Of the 284 individual adult *I. scapularis* tested for presence of *B. burgdorferi* using a multiplex real-time PCR assay (Hojgaard et al. 2014), 35 (12.3%) were positive (Table 4.7.). The percent of tested ticks that were positive varied across the parks from a low of zero (where only one tick was tested) to a high of 17.7% (Table 4.7.). An index was created to assess density of infected adult ticks; this index was calculated as the number of positive ticks per 100 deer culled. Values for this index ranged from a low of 8.3 to a high of 27.9 across parks (Table 4.7.).

Reported human Lyme disease among a subset of Howard County households

A total of 21,370 homes were located within half-mile radius buffers around 15 selected parks. The number of homes in an individual park buffer ranged from 159 to 5,040 homes. Overall, at least one reported case of Lyme disease during 2001-2011 occurred in 451 (2.1%) of these homes. Cumulative incidence of Lyme disease during the study period varied between 0.58% of homes to 5.03% of homes around a park (Table 4.6.).

Association between deer density and blacklegged tick abundance

Univariate associations

The association between tick abundance on deer during October and median deer density was examined. In addition to median density as a continuous variable (the primary explanatory variable of interest), the seven parks with hunting data were also grouped into a three-level deer density variable based on median density (≤ 24, 25 - 62, > 63 deer per mi^2). A visual depiction of the crude association between mean tick abundance and median deer density is shown in Figures 4.3 and 4.4. Based upon these figures, an inverse association between tick abundance and deer density was expected.

Three different regression model types were explored: Poisson, overdispersed Poisson, and negative binomial models. Variance in tick counts during October (σ^2=11.9) exceeded the mean (μ=1.4), thus a pure Poisson model was not considered further.

Table 4.8 displays relative risks (RR) for two deer density variable specifications and different model types. In the crude overdispersed Poisson model, median deer density as a continuous variable was inversely associated with tick abundance on deer during October (RR

per 10 deer per mi^2 increase in density: 0.92, 95% CI: 0.84-0.99, p=0.033). Upon examination of the three-level class variable for median deer density, the lower two categories of deer density (density ≤ 24 deer per mi^2 and 25 - 62 deer per mi^2) did not differ with respect to counts of ticks on deer. In contrast, the highest level of deer density was borderline inversely associated with tick abundance on deer compared to the lowest level of deer density (RR: 0.54, 95% CI: 0.28-1.03, p=0.063). Although the negative binomial model is generally more flexible in handling overdispersion than a Poisson model, the Poisson model has a simpler variance structure (Gardner et al. 1995). Given that the standard errors did not differ substantially, the simpler overdispersed Poisson model was selected for a multivariable model.

Multivariable model development

In an adjusted overdispersed Poisson model that accounted for sex and age of deer and year of cull, the inverse association between median deer density and tick abundance on deer was stronger when compared to unadjusted models, regardless of how deer density was specified (RR per 10 deer per mi^2: 0.87, 95% CI: 0.81-0.94, p<0.001). Additional adjustment for total land area at each location assessed to determine density did not change the strength or direction of the association.

The need to account for non-independence of deer within parks was examined by comparing results to those from a generalized estimating equations (GEE) framework using an exchangeable covariance structure. The working correlation matrix value of -0.0022 and a minimal change in QIC value between the GEE with exchangeable correlation (QIC=207.8) and assuming independence between observations (QIC=208.8) indicated no need to account for

clustering. Additionally, the magnitude of the relative risk estimates for both deer density variable specifications did not change substantively with the GEE framework (Table 4.8.). The overdispersed Poisson model without GEE had larger (i.e., more conservative) standard errors and thus was considered the final model.

Seven observations displayed high deviance residuals; these observations had very high tick counts. To examine the impact of these outlying values, a multinomial logistic model (outcome values of 0, 1, ≥ 2 ticks) and a Poisson model after truncating higher end tick counts at 10 were constructed. The findings of the highest deer density associated with the lowest tick counts were robust to specification of the outcome variable. The association of tick counts on deer and deer density was also assessed using alternate deer density specifications (mean and variance), and alternate tick abundance metrics. Mean density was highly correlated with median; findings were similar. Parks with the highest average density and lowest tick abundance also had the highest variance in deer density estimates over the study period. This may indicate that the higher densities reflect more instability in how deer use the environment and their accurate quantification. Results from these supplemental analyses are presented in Appendices 3.1.-3.3.

Association of deer density with risk of human Lyme disease

Univariate associations

A scatterplot of cumulative incidence of Lyme disease around parks by quartile of median deer density is shown in Figure 4.5. Visually, parks with higher average deer density did not display higher incidence of Lyme disease in surrounding buffers as was hypothesized. In

univariate analyses, with only a fixed effect for quartile of median deer density, there was an unexpected trend toward an inverse association between median deer density and risk of human Lyme disease. Specifically, the two highest quartiles of deer density were associated with lower risk of human disease compared to the lowest quartile of deer density (Q3 OR: 0.37, 95% CI: 0.23-0.61; Q4 OR: 0.64, 95% CI: 0.52-0.79). However, despite a highly statistically significant p-value for trend across quartiles, there was not a clear linear or threshold association with reduced human risk across the quartiles; the third quartile of deer density displayed a stronger inverse association than the highest quartile.

Multivariable model development

Univariate associations of potential covariates with human disease are shown in Table 4.9. Potential covariates assessed at the park and buffer could not be included in multivariable models because they were collinear with deer density assessed at the park (i.e., all houses surrounding a park had the same value for these variables). The final adjusted model contained class variables for land use and land cover type at the household location and a dichotomous variable for median home value in the census block group. After adjustment for confounding factors, there was not a clear association between deer density and human Lyme disease risk, although the trend toward lower human risk in areas with higher deer density remained (Table 4.10.). Specifically, the first and second quartiles were not statistically different from one another; the inverse association between the highest deer density level and human disease risk was attenuated compared to that of the third quartile of deer density (Q3 OR: 0.49, 95% CI: 0.30-0.80; Q4 OR: 0.77, 95% CI: 0.60-1.01).

In an adjusted GEE model, the exchangeable working correlation matrix was very small (0.00031). QIC values for the exchangeable structure (QIC= 4306.8) and for an independent structure (QIC=4303.6) were highly similar, indicating minimal need to account for correlation among homes given the covariates included. Not surprisingly, accounting for potential clustering did not substantially affect statistical significance or interpretation; the first and second quartiles did not differ from one another, the third quartile of deer density was associated with reduced risk of Lyme disease compared to the lowest quartile of density (OR: 0.49, 95% CI: 0.32-0.75, p=0.001), and the highest level of deer density was not significantly different than the lowest two quartiles with respect to human disease risk (OR: 0.75, 95% CI: 0.52-1.09, p=0.134; Table 4.10.).

The final adjusted logistic model without a GEE structure was considered the final model. This model displayed moderate goodness-of-fit. Specifically, the Deviance test indicated good fit (p=0.29), although the Pearson test was highly significant (p <0.001); the Hosmer and Lemeshow goodness-of-fit test indicated moderate fit (p=0.32). No deviance residuals were higher than four; although there were a few outlying observations, all remained in the model.

Association between deer density and human Lyme disease risk was also examined using other deer density metrics (mean and variance), and findings were similar. Areas with highest deer density regardless of average metric and with the highest variance were associated with lowest risk of disease (Appendix 3.4.).

Association of tick infestation on deer with risk of human Lyme disease

Univariate associations

The association between mean tick infestation on hunted deer and human Lyme disease was assessed in the buffer area surrounding parks with hunting data (seven parks, n=12,881 households). Mean number of ticks on deer during October as a continuous variable was not associated with risk of human disease in univariate analysis (OR per one unit increase in mean tick count: 0.99, 95% CI: 0.75-1.31; p=0.963), nor was the dichotomized mean number of ticks during October (> 1.37 vs. ≤ 1.37 ticks) (OR: 1.14, 95% CI: 0.90-1.44, p=0.287; Table 4.11.).

Additionally, two alternative measures of tick abundance, variance and maximum, were investigated for their association with human Lyme disease. Variance in number of ticks on deer during October, as a continuous variable, was not associated with human risk (OR per one unit increase in variance: 1.0, 95% CI: 0.98-1.01, p=0.851), nor was variance as a dichotomous variable (variance > 10 vs. ≤ 10) (OR: 1.21, 95% CI: 0.94-1.56, p=0.144). Maximum number of ticks found on a single deer in a park, without restriction on months, as a continuous variable, was associated with a slightly increased risk of human disease in the surrounding area (OR per single tick increase in maximum ticks on deer: 1.01, 95% CI: 1.00-1.02, p=0.041; OR per 10 tick increase in maximum ticks on deer: 1.12, 95% CI: 1.00-1.24; Table 4.11.). When maximum number was dichotomized (< 25 ticks, ≥ 25 ticks), the univariate association with human risk was stronger (OR for ≥ 25 ticks: 1.41, 95% CI: 1.10-1.81, p=0.007; Table 4.11.).

Multivariable model development

In a model of mean number of ticks on deer as a dichotomous variable, land use class at the household was the only confounding variable. In a minimally adjusted model with only this covariate, mean number of ticks on deer was associated with elevated human Lyme disease risk (OR for >1.37 ticks vs. ≤ 1.37 ticks): 1.29, 95% CI: 1.01-1.63; p=0.039; Table 4.11.). A fully adjusted model was also created that included land use and land cover at the household location and a dichotomous variable for median home value; this model was constructed to increase comparability with the deer density and human Lyme disease model. In the fully adjusted model, the association between mean number of ticks on deer and human Lyme disease risk diminished (OR for > 1.37 ticks vs. ≤ 1.37 ticks): 1.15, 95% CI: 0.90-1.48; p=0.257). Mean number of ticks specified as a continuous variable was not associated with human risk in multivariable models (Table 4.11.).

The fully adjusted model fit the data well. Specifically, the Deviance and Pearson tests were highly non-significant (p=0.712 and p=0.395, respectively); the Hosmer and Lemeshow goodness-of-fit test had a p-value of 0.949. Deviance residuals were nearly all below three.

Upon comparison with a fully adjusted model constructed with an exchangeable covariance structure, the working correlation matrix for the model with mean tick infestation as a dichotomous variable was slightly larger than in other models above (0.001). Moreover, QICs varied more between the exchangeable and independent covariance structures (QIC exchangeable=2,757; QIC independent= 2,737). Although interpretation did not differ, more

unexplained correlation among homes was evident in this analysis; the GEE model was considered the final model.

In an exploratory multivariable analysis of maximum number of ticks on deer in a park as the explanatory variable, a minimally adjusted model as above (including land use class at household) was associated with elevated risk as both a continuous and dichotomous variable. The association was attenuated in the fully adjusted model (OR for > 25 vs. ≤ 25 ticks: 1.35, 95% CI: 0.96-1.91; p=0.086), but a trend toward increased risk persisted (Table 4.11). However, in a GEE model no association remained (Table 4.11). Additionally, median deer density as a continuous variable was added to a fully adjusted model with dichotomized maximum ticks; neither the tick variable, nor deer density, were associated with human disease risk after adjusting for household land use and land cover.

Finally, a model was constructed to assess the association between infestation of deer by infected adult blacklegged ticks from the 2012-2013 deer management season and risk of Lyme disease during the entire study period; there was no detectable association (Appendix 3.5.).

DISCUSSION

This study examined the associations of deer density, tick infestation, and human Lyme disease risk on a sub-county scale. In this analysis, deer density was not clearly and strongly associated with human Lyme disease risk. Nevertheless, there was evidence of a trend toward higher disease risk in areas with lower deer density but higher tick infestation on deer.

Deer density varied substantially across time and space, and the difficulty of accurate quantification may have affected study findings. However, results from this study suggest that moderate difference in deer density across space in an endemic county is not an obvious cause of spatial variation in risk of human Lyme disease. Intervention studies assessing the impact of deer population reduction on blacklegged tick abundance suggest that, in some circumstances, deer may need to be nearly absent from an area to affect tick reproduction (Wilson et al. 1984; Wilson et al. 1988; Rand et al. 2004). However, the mechanism through which deer could contribute to human Lyme disease risk may occur on a very fine spatial scale. Difference in utilization of specific properties by deer was not assessed, but may still contribute to spatial differences in risk to humans on their respective properties. Three studies have found self-reported evidence of deer utilizing the residential property to be associated with increased Lyme disease risk (Lastavica et al. 1989; Orloski et al. 1998; Smith et al. 2001).

Tick infestation on deer varied across parks, and may be a more meaningful measure of human risk than deer density. Several studies conducted at larger spatial scales have identified positive associations between tick abundance on deer and human Lyme disease, but this study was the first known to be conducted within a single county (Wallis et al. 1978; Schulze et al. 1984; Daniels et al. 1993a; Kitron and Kazmierczak 1997; Raizman et al. 2012). The most appropriate measure of tick infestation on deer as it may relate to the enzootic cycle and quantification of tick population abundance across space is unknown; here, we explored mean, maximum and variance. Low statistical power limited the ability to detect minimal differences in overall infestation across space given the low tick counts, collected from only the right ear of hunted deer. If more of each deer had been examined and more differentiation between high

and low infestations was possible, a stronger association between tick infestation and human disease risk may have been detected. Moreover, male deer had twice as many ticks as female deer, an observation that has been noted previously (Main et al. 1981; French et al. 1992; Amerasinghe et al. 1993; Cortinas and Kitron 2006). These consistent findings underscore that the association between deer and human disease risk is complex; simply managing a deer population within a single inland jurisdiction will likely not impact human Lyme disease risk.

In this project, residence around parks with more red and white oak forest and larger forest patches (despite the same relative amount of forest) was associated with increased human risk in univariate analysis. Teasing apart the relative contributions of these factors in concert with deer density, tick infestation on deer and human disease risk was not possible in this project. Likewise, the relative importance of deer density, tick infestation on deer, and land use and land cover with respect to human Lyme disease could not be distinguished; all homes around a park had the same value for deer density, tick abundance, proportion of park and buffer with a given land use or land cover types and could not all be included in a statistical model together. Prospective data collection that includes more individual, household-specific and census block group level data may allow these types of analyses in the future.

Limitations

This study was subject to several limitations. Existing data on tick infestation on deer (counts on the right ear of hunted deer) were utilized, which may have limited statistical power to detect differences in average tick infestation across parks in Howard County. Counts of ticks on deer could have been subject to measurement error not only because of the small portion of

the deer examined, but also because of error in tick detection and quantification. This measurement error would likely be more pronounced in the higher tick counts, where quick counting, losing track of numbers, or not counting as carefully after reaching a certain number of ticks, were all possible. In fact, we witnessed this type of discrepancy during the 2012-2013 management season between recorded tick counts per deer and number of actual ticks removed and placed in vials. Ticks on highly infested deer were undercounted compared to the number ultimately identified in vials and tested for *B. burgdorferi*. This underestimation was likely non-differential across parks; tick counts were performed by one of two designated people during the entire study period. Undercounting of higher numbers of ticks would serve to bias findings toward the null. Species was determined for all ticks collected during the 2012-2013 season, and only one was not *I. scapularis*. Although all historically counted ticks were probably not *I. scapularis*, species determination from the most recent year provided some confidence that most were likely the species of interest.

Misclassification of the deer density variable was likely, and was probably a major source of bias in the deer density results. Density was assessed at specific parks and natural areas that may not be representative of the rest of Howard County. Although the aerial surveys occurred only once per year, and deer are a dynamic population in time and space, attempts were made to minimize bias. A categorized cumulative variable was used to minimize the impact the expected misclassification may have had on the findings. Nevertheless, median density over time as assessed here was not clearly associated with risk of Lyme disease. Yet, a biologically plausible inverse association was detected between tick infestation on deer and deer density using this method of deer density measurement. Increase in tick abundance on

deer in concert with lower deer density was detected temporally in intervention studies (Deblinger et al. 1993; Rand et al. 2004). Presumably, in areas with lower deer density, the remaining deer host more ticks. Quantification of deer pellet counts in each park prior to the beginning of each hunting season may have been a more accurate and cheaper, albeit much more labor-intensive, alternative to assess population density. Although use of a cumulative measure compared to a time-lagged density variable may have resulted in loss of information, the issue of measurement error outweighed potential loss of information. Moreover, the complexity of assigning a time lag to control houses that did not have inherent associated date of "outcome" was beyond the scope of this project.

Furthermore, assumption that all homes within a half mile of parks are equally exposed to the deer density of that park is likely inaccurate. Performing similar analysis using increasingly smaller buffer areas around parks could prove informative; however, small numbers of homes around more rural parks may limit interpretability of more restricted buffer sizes. Residual confounding due to differences in land use and land cover in the buffer areas was possible; nevertheless, controlling for land use and land cover at each household location in the buffer areas controlled for some of these differences in acceptability of the buffer area to deer.

Misclassification of households according to Lyme disease status was possible, as was discussed in earlier chapters. Any potential for differential diagnosis and reporting of Lyme disease according to socio-economic status discussed in earlier studies, although still possible,

may have been minimized by using a restricted dataset of homes that were all within close proximity to natural areas, generally indicative of similar socio-economic status.

In these analyses, location of residence was assumed to be the location of exposure. Lyme disease is widely assumed to be acquired peridomestically, and thus use of geocoded residence is an established surrogate for site of exposure (Maupin et al. 1991; Orloski et al. 1998; Eisen and Eisen 2007). In this analysis the same deer density and tick infestation on deer was assigned to each home around a park. Consequently, the potential for misclassification of exposure location was minimized; the assumed exposure location was not restricted to one property only but therefore included the nearby park and the entire buffer area around the park.

Deer density and tick infestation on deer are separated from human disease risk by two years of blacklegged tick life cycle, and factors including weather and reservoir host abundance, all of which likely play the biggest role in determining the density of infected nymphal blacklegged ticks and which were unmeasured in this analysis. Furthermore, adult tick infestation on deer may not directly correlate with abundance of nymphal ticks two years later, as multiple factors affect tick survival. These intermediate factors, as well as human interaction with the environment all affect human risk. Furthermore, as many of these factors likely differ across Lyme disease-endemic areas, generalizability of findings in this study beyond Howard County may be limited.

Strengths

This was the first study to examine the associations of deer density, tick infestation on deer, and human Lyme disease risk on a sub-county scale. Use of data collected over several years increased power and reliability of findings. Additionally, use of data from a single county minimized the differences in exposure and outcome assessment that may occur across jurisdictions.

Linear distance from a home to a wooded area or park could be associated with Lyme disease risk but was unmeasured in Project 2. If access to these types of areas affects risk, the restriction of analyses in Project 3 to homes within one-half mile of wooded natural areas minimized potential confounding by that unmeasured variable. Interestingly, the same household land use and land cover characteristics were associated with elevated risk in this restricted dataset as were in Project 2.

Although intervention studies to assess effect of deer density on human Lyme disease risk are needed, this study used observational data collected by a community deer management program to assess the role that moderate differences in deer density across space may play in risk of human Lyme disease. Existing data were used to assess how a typical community deer management program's actions may affect risk of Lyme disease in their community, rather than determining an association from artificial experimental conditions that lack applicability to real-world prevention programs. These findings suggest that while deer overpopulation is problem for numerous reasons, including risk of motor-vehicle accidents and

crop damage, moderate community deer management activities in inland communities may have minimal impact on human Lyme disease risk in already endemic areas.

Conclusions

Deer reduction has been proposed as a Lyme disease prevention measure, yet the association between deer density and human Lyme disease risk is unknown. This study provides some evidence that differences in average deer density over time may not be tied to sub-county spatial variation in risk of Lyme disease in endemic areas. Here, higher burden of ticks per deer was demonstrated in areas with lower deer density, but that finding could result from other land cover differences not assessed in this project. Nevertheless, it is entirely possible that the association between deer density and human risk was muted by a higher number of ticks per deer where deer are less plentiful. Future research into the role of spatial differences in deer density and Lyme disease risk should use one or more alternate means of assessing density, and attempt to obtain ticks from a larger portion of individual deer than just the right ear when assessing abundance.

TABLES

Table 4.1. Deer density* across parks by year, Howard County, Maryland

Year	**No. parks assessed**	**Total deer density** (total deer/ total area)	**Mean deer density** (per park)	**Median deer density** (per park)	**Range in deer density** (across parks)
2001	10	82.8	92.7	51.5	17 - 448
2002	9	82.3	89.7	64.0	17 - 285
2003	10	94.3	102.8	111.5	22 - 178
2004	11	83.3	103.1	61.5	17.7 - 275
2005	7	34.3	30.8	26.0	14 - 48
2006	12	78.4	124.2	64.0	22 - 530
2007	14	47.8	76.7	40.5	0 - 540
2008	15	79.1	89.1	56.0	18 - 425
2009	14	51.2	74.9	44.0	17.2 - 450
2010	14	35.2	78.1	26.6	1 - 535
2011	10	9.9	18.8	18.9	0 - 41.4

*Assessed as deer/mi^2

Table 4.2. Deer density* estimated in individual parks, Howard County, Maryland, 2001-2011

Park	Size assessed (mi^2)	No. years	No. years hunted	Mean density	Median density	Range in density	Variance in density	Quartile of median deer density
ARP	0.17	11	10	39.3	29.4	12 - 75	511.5	1
DML	0.69	6	0	36.2	37.8	2 - 65	571.6	1
DFP	0.50	11	13	41.0	34.0	14 - 96	677.8	1
MPE	1.75	10	15	25.4	24.0	14 – 53	125.3	1
BBP	1.80	8	0	45.1	38.2	22 – 99	665.8	2
SAV	5.35	5	0	42.1	43.6	1 – 117	2,255.4	2
SMP	0.29	8	7	46.6	39.7	17- 107	809.3	2
WRP	0.29	7	0	50.4	45.0	10 - 99	1,373.8	2
HRP	0.28	5	7	48.7	47.0	0 - 100	1,277.5	3
WOO	0.36	11	0	106.3	55.6	14 - 275	8,629.5	3
WFP	1.09	11	3	75.8	64.0	21 - 161	2,706.5	3
BEL	0.20	5	0	496	530.0	425 - 540	2,942.5	4
BLA	0.50	10	10	133.2	104.0	38 - 448	14,590.4	4
MBP	0.27	8	0	76.2	76.5	0 - 161	3,905.7	4
RBP	0.63	9	0	108.1	67.0	17 - 285	7,060.3	4

*Assessed as deer/mi^2

Table 4.3. Deer culled and mean number of ticks identified on right ear of hunted deer

Year*	No. parks hunted	Total no. deer culled	Mean no. deer culled per park	Range no. culled per park	Total ticks counted	Mean no. ticks per deer
1997	1	50	50	n/a	9	0.18
1998	n/a	n/a	n/a	n/a	n/a	n/a
1999	1	134	134	n/a	253	1.89
2000	2	256	128	96 - 160	159	0.62
2001	2	164	82	56 - 108	67	0.41
2002	2	90	45	30 - 60	165	1.83
2003	4	199	49.8	32 - 60	146	0.73
2004	4	220	55	39 - 78	105	0.48
2005	4	177	44.3	27 - 80	79	0.45
2006	7	130	18.6	5 - 43	115	0.88
2007	6	188	31.3	4 - 90	124	0.66
2008	7	188	26.9	1 - 56	164	0.87
2009	9	237	26.3	4 - 72	259	1.09
2010	7	189	27	9 - 51	182	0.96
2011	7	181	25.9	2 - 80	259	1.43
2012	7	200	28.6	9 - 68	267	1.34

*Year classified according to beginning of deer management season (fall), but associated data collection continued through following winter. No deer management activities took place during the management season that began in 1998.

Table 4.4. Cumulative number of deer culled and tick abundance on deer by park, Howard County, Maryland

Park	Quartile median deer density	Total no. deer culled	Total hunting days	Deer culled per hunting day	Mean no. ticks per deer
ARP	1	274	49	5.6	1.0
DML	1	0	0		
DFP	1	479	93	5.2	1.0
MPE	1	1,149	147	7.8	1.1
BBP	2	0	0		
SAV	2	0	0		
SMP	2	123	26	4.7	0.7
WRP	2	0	0		
HRP	3	60	17	3.5	1.1
WOO	3	0	0		
WFP	3	57	7	8.1	0.3
BEL	4	0	0		
BLA	4	345	69	5	0.4
MBP	4	0	0		
RBP	4	0	0		

Table 4.5. Deer culled during October only and mean number of ticks identified on right ear of hunted deer, by year, Howard County

Year	No. parks hunted	Total deer culled	Total ticks counted	Mean no. ticks per deer
1997	0	-	-	-
1998	0	-	-	-
1999	1	102	185	1.88
2000	2	73	58	0.82
2001	2	59	36	0.62
2002	-	-	-	-
2003	4	56	78	1.42
2004	4	59	51	0.88
2005	4	49	16	0.33
2006	5	37	45	1.22
2007	5	62	47	0.78
2008	5	75	107	1.43
2009	6	84	128	1.54
2010	3	25	33	1.38
2011	5	46	152	3.38
2012	4	32	83	2.59

Table 4.6. Deer density, deer culled, tick abundance on deer in October, and cumulative incidence of Lyme disease in buffer areas, by park, Howard County, Maryland

Park	**Quartile median deer density**	**Total deer culled** (Oct)	**Total hunting days** (Oct)	**Deer culled per hunting day** (Oct)	**Mean no. ticks per deer** (Oct)	**Variance in tick count** (Oct)	**No. houses in buffer**	**Cumulative Lyme disease incidence in buffer** (%)*
ARP	1	65	11	5.9	2.63	41.52	515	1.36
DML	1	-	-	-	-		740	3.51
DFP	1	124	22	5.6	1.37	15.47	2,190	2.15
MPE	1	358	36	9.9	1.40	10.15	5,040	2.82
BBP	2	-	-	-	-		576	5.03
SAV	2	-	-	-	-		1,590	1.07
SMP	2	23	5	4.6	0.87	1.76	320	4.69
WRP	2	-	-	-	-		208	3.85
HRP	3	22	5	4.4	1.41	4.35	840	0.60
WOO	3	-	-	-	-		799	1.38
WFP	3	5	1	5.0	1.60	12.80	159	1.26
BEL	4	-	-	-	-		1,031	0.58
BLA	4	111	19	5.8	0.73	2.11	3,817	1.86
MBP	4	-	-	-	-		1,632	1.84
RBP	4	-	-	-	-		1,913	1.83

*Percent of homes with reported Lyme disease during 2001-2011

Table 4.7. *Borrelia burgdorferi* infection prevalence among adult blacklegged ticks, Fall 2012-Winter 2013

Park	**Deer culled** (N)	**Deer infested*** n (%)	**Ticks tested** (N)	**Positive ticks**** n (%)	**Density of infected adults** (positive ticks per 100 deer culled)	**Cumulative Lyme disease incidence in buffer** (%)†
ARP	33	7 (21)	24	3 (12.5)	9.1	1.36
BLA	9	1 (11)	1	0 (0)		1.86
DFP	41	14 (34)	72	8 (11.1)	19.5	2.15
HRP	11	0 (0)	0	n/a		0.6
MPE	68	23 (34)	157	19 (12.1)	27.9	2.82
SMP	24	3 (13)	13	2 (15.4)	8.3	4.69
WFP	14	3 (21)	17	3 (17.7)	21.4	1.26

*Infested with at least one tick on the right ear

**Positive by real-time PCR assay designed to detect presence of *B. burgdorferi* in ticks

†Percent of homes with reported Lyme disease during 2001-2011

Table 4.8. Unadjusted, adjusted, and generalized estimating equations (GEE) models of association of tick counts on deer and deer density in parks, Howard County, Maryland*

* Relative risk (RR); 95% confidence intervals (CI) presented; reference group indicated for class variables (ref)

Deer density variable	**Variable specification**	**Crude overdispersed Poisson model****			**Crude negative binomial model**			**Adjusted overdispersed Poisson model**†**			**Adjusted Poisson with GEE†‡**		
		RR	*95% CI*	*p-val*	*RR*	*95% CI*	*p-val*	*RR*	*95% CI*	*p-val*	*RR*	*95% CI*	*p-val*
Median	10 deer per mi^2	0.92	0.84-0.99	0.033	0.91	0.86-0.97	0.003	0.87	0.81-0.94	<0.001	0.88	0.82-0.94	<0.001
Low	≤ 24 deer per mi^2	*ref*			*ref*			*ref*			*ref*		
Medium	25 - 62 deer per mi^2	1.23	0.85-1.77	0.272	1.23	0.88-1.72	0.231	1.30	0.91-1.87	0.153	1.26	0.98-1.63	0.068
High	≥ 63 deer per mi^2	0.54	0.28-1.03	0.063	0.54	0.34-0.85	0.009	0.45	0.24-0.82	0.009	0.44	0.34-0.57	<0.001

**Overdispersion in Poisson model scaled by square root of Pearson's Chi-square/degrees of freedom

†Models adjusted for sex and age of deer and year culled

‡ GEE implemented with exchangeable correlation structure

Table 4.9. Univariate associations of possible covariates with risk of Lyme disease among homes surrounding parks, Howard County, Maryland

Variable	**Variable specification**	**N**	**Houses with Lyme disease** n (%)	**Odds ratio†**	**95% CI†**	**p-val**
Land use*	Urban/high-density development	5,944	73 (1.2)	*ref*		
	Herbaceous/ agriculture	614	5 (0.8)	0.66	0.27-1.66	0.382
	Forest	922	17 (1.9)	1.51	0.89-2.57	0.129
	Medium-density development	10,629	259 (2.4)	2.01	1.55-2.61	<0.001
	Low-density development	3,250	97 (3.0)	2.47	1.82-3.36	<0.001
Land cover*	Urban	10,638	194 (1.8)	*ref*		
	Red-white oak forest	2,288	86 (3.8)	2.10	1.63-2.72	<0.001
	Other forest type	1,559	35 (2.3)	1.24	0.86-1.78	0.253
	Herbaceous/ agriculture	6,869	136 (2.0)	1.09	0.87-1.36	0.458
Four poster tick control **	No	17,553	380 (2.2)	*ref*		
	Yes	3,817	17 (1.9)	0.86	0.66-1.11	0.236
Median home value ***	Q1	4,902	55 (1.1)	*ref*		
	Q2	5,445	107 (2.0)	1.76	1.27-2.44	<0.001
	Q3	5,576	115 (2.1)	1.85	1.34-2.56	<0.001
	Q4	5,547	174 (3.2)	2.90	2.14-3.94	<0.001
Median household income ***	Q1	5,260	85 (1.6)	*ref*		
	Q2	4,808	75 (1.6)	0.96	0.70-1.32	0.807
	Q3	5,901	126 (2.1)	1.32	1.00-1.75	0.047
	Q4	5,401	165 (3.05)	1.91	1.47-2.49	<0.001
Per capita income ***	Q1	4,864	67 (1.4)	*ref*		
	Q2	5,720	104 (1.8)	1.32	0.97-1.80	0.078
	Q3	5,155	107 (2.1)	1.51	1.11-2.06	0.008
	Q4	5,631	173 (3.1)	2.26	1.70-3.00	<0.001
Population density ***	Q1	4,979	111 (2.2)	*ref*		
	Q2	5,244	145 (2.8)	1.24	0.97-1.60	0.087

	Q3	5,226	80 (1.5)	0.68	0.51-0.91	0.009
	Q4	5,921	115 (1.9)	0.87	0.67-1.13	0.287
Population growth 2000-2009 ***	Q1	5,782	119 (2.1)	*ref*		
	Q2	5,580	127 (2.3)	1.11	0.86-1.42	0.435
	Q3	5,180	99 (1.9)	0.93	0.71-1.21	0.577
	Q4	4,828	106 (2.2)	1.07	0.82-1.39	0.635
Proportion low-density development **	Q1	6,438	94 (1.5)	*ref*		
	Q2	2,631	37 (1.4)	0.96	0.66-1.41	0.840
	Q3	4,843	108 (2.2)	1.54	1.16-2.03	0.003
	Q4	7,458	212 (2.8)	1.97	1.54-2.52	<0.001
Proportion red-white oak forest **	Q1	3,178	43 (1.4)	*ref*		
	Q2	7,413	114 (1.5)	1.14	0.80-1.62	0.470
	Q3	5,163	123 (2.3)	1.78	1.25-2.52	0.001
	Q4	5,616	171 (3.0)	2.29	1.63-3.21	<0.001
Proportion forest **	Q1	3,154	56 (1.8)	*ref*		
	Q2	5,424	106 (2.0)	1.10	0.80-1.53	0.556
	Q3	7,709	206 (2.7)	1.52	1.13-2.05	0.006
	Q4	5,083	83 (1.6)	0.92	0.65-1.29	0.632
Mean forest patch size **	Q1	5,086	89 (1.8)	*ref*		
	Q2	6,245	113 (1.8)	1.03	0.78-1.37	0.815
	Q3	4,999	107 (2.1)	1.23	0.93-1.63	0.155
	Q4	5,040	142 (2.8)	1.63	1.24-2.13	<0.001
Forest edge density **	Q1	5,780	168 (2.9)	*ref*		
	Q2	5,922	95 (1.6)	0.55	0.42-0.70	<0.001
	Q3	4,582	99 (2.2)	0.74	0.57-0.95	0.018
	Q4	5,086	89 (1.8)	0.60	0.46-0.77	<0.001

†Variables divided into quartiles, reference group indicated (ref); 95% CI=95% confidence interval

* Variable assessed at house point location

**Variable assessed at park/buffer. Proportions were calculated as proportion of total park/buffer area of that class

***Variable assessed at census block group

Table 4.10. Unadjusted, adjusted, and adjusted generalized estimating equations (GEE) logistic regression models of deer density and human Lyme disease risk, Howard County, Maryland, 2001-2011*

Variable	Variable specification	Unadjusted			Adjusted**			Adjusted + GEE†		
		OR	*95% CI*	*p-value*	*aOR*	*95% CI*	*p-value*	*aOR*	*95% CI*	*p-value*
Deer density	Q1 (< 38 deer per mi^2)	*ref*			*ref*			*ref*		
	Q2 (38 - 45 deer per mi^2)	0.98	0.74-1.29	0.876	1.17	0.86-1.59	0.329	1.25	0.75-2.07	0.392
	Q3 (46 - 56 deer per mi^2)	0.37	0.23-0.61	<0.001	0.49	0.30-0.80	0.004	0.49	0.32-0.75	0.001
	Q4 (> 56 deer per mi^2)	0.64	0.52-0.79	<0.001	0.77	0.60-1.01	0.055	0.75	0.52-1.09	0.134
Land use class	Low-density development				1.79	1.26-2.54	0.001	1.72	1.36-2.18	<0.001
	Medium-density development				1.78	1.36-2.38	<0.001	1.79	1.34-2.39	<0.001
	Herbaceous/agriculture				0.57	0.23-1.44	0.236	0.56	0.21-1.51	0.253
	Forest				1.10	0.63-1.90	0.744	1.08	0.62-1.87	0.789
	Urban/high-density development				*ref*			*ref*		
Land cover class	Herbaceous/agriculture				0.85	0.66-1.10	0.223	0.85	0.71-1.32	0.070
	Red-white oak forest				1.60	1.20-2.10	0.001	1.55	1.22-1.98	<0.001
	Other forest				1.01	0.69-1.48	0.953	1.01	0.78-1.32	0.941
	Urban				*ref*			*ref*		
Median home value	High				1.36	1.02-1.81	0.036	1.34	0.91-1.99	0.140
	Low				*ref*			*ref*		

*Odds ratios (OR), adjusted odds ratios (aOR), 95% confidence intervals (95% CI), and p-values presented; reference group for class variables indicated (ref).

**Adjusted for land use type at house, land cover type at house and dichotomous median home value at census block group

†GEE model constructed with exchangeable covariance structure

Table 4.11. Unadjusted, adjusted, and adjusted generalized estimating equation (GEE) logistic regression models of tick abundance on deer during October and human Lyme disease risk, Howard County, Maryland*

Variable	Variable specification	Unadjusted			Minimally adjusted**			Fully adjusted†			Adjusted + GEE‡		
		OR	95% CI	p-val	aOR	95% CI	p-val	aOR	95% CI	p-val	aOR	95% CI	p-val
Mean ticks on deer (Oct)	*per one unit increase*	0.99	0.75-1.31	0.963	1.20	0.87-1.63	0.278	1.00	0.68-1.47	0.996	0.88	0.44-1.78	0.725
	Low (≤1.37 ticks)	*ref*			*ref*			*ref*			*ref*		
	High (>1.37 ticks)	1.14	0.90-1.44	0.287	1.29	1.01-1.63	0.039	1.15	0.90-1.48	0.257	0.93	0.49-1.74	0.815
Maximum ticks on deer	*per one unit increase*	1.01	1.00-1.02	0.041	1.01	1.00-1.02	0.021	1.01	0.99-1.02	0.285	1.01	0.99-1.08	0.410
	Low (<25 ticks)	*ref*			*ref*			*ref*			*Ref*		
	High (≥25 ticks)	1.41	1.10-1.81	0.007	1.46	1.14-1.88	0.003	1.35	0.96-1.91	0.086	1.29	0.68-2.42	0.427

*Odds ratios (OR), adjusted odds ratios (aOR), 95% confidence intervals (95% CI), and p-values presented; reference group for class variables indicated (ref).
**Adjusted for land use
†Adjusted for land use and land cover at household location and dichotomous median home value at census block group
‡Used fully adjusted model; GEE specified with an exchangeable covariance structure

FIGURES

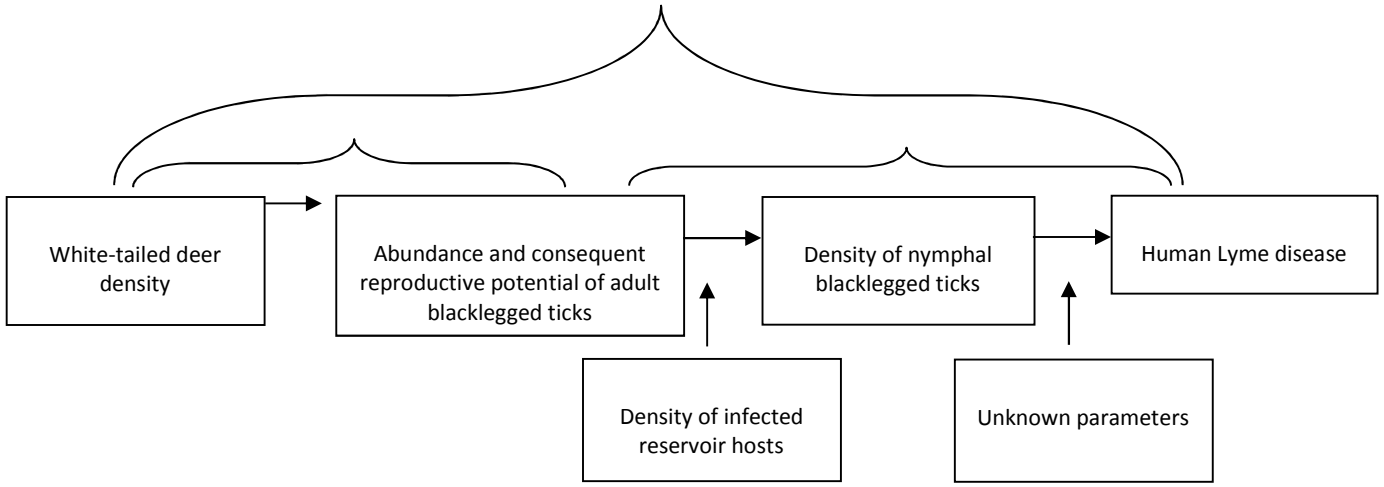


Figure 4.1. Simplified mechanism through which deer act on human Lyme disease with representation of associations evaluated in this study.

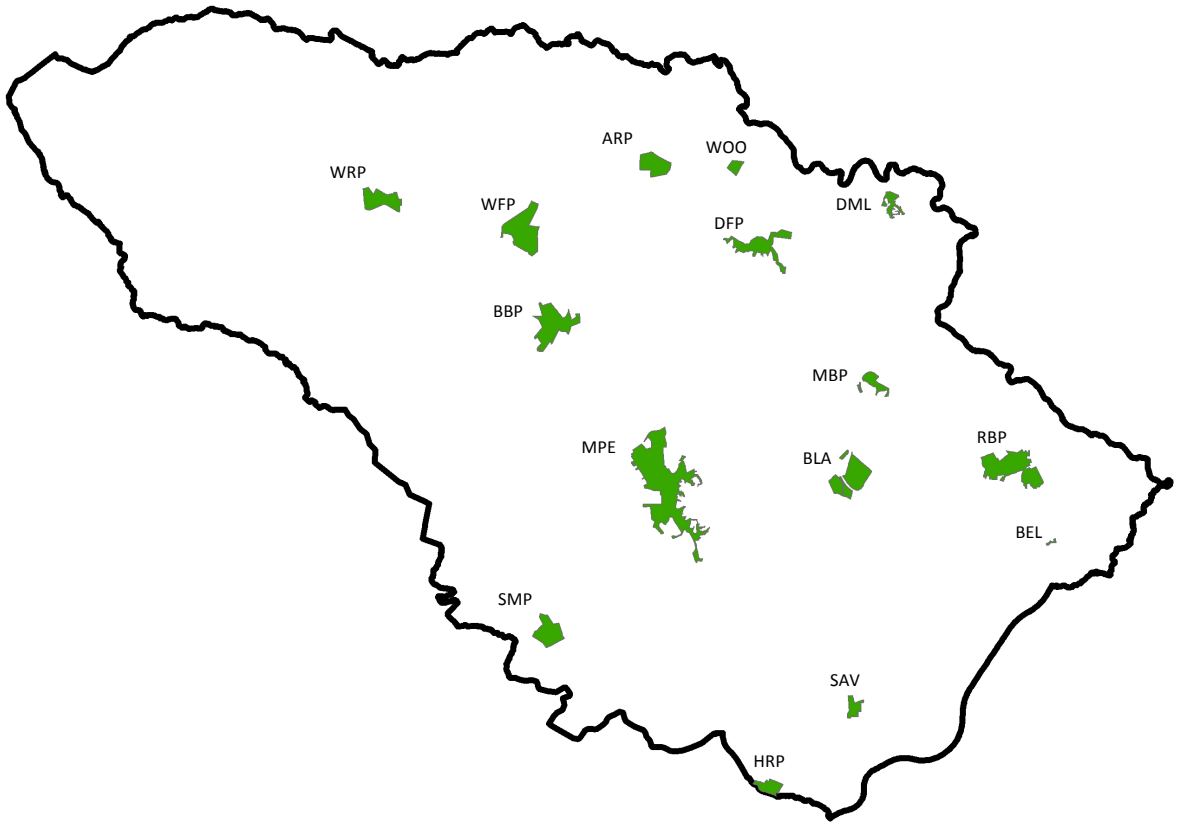


Figure 4.2. Map of Howard County with parks used in analysis shown in purple

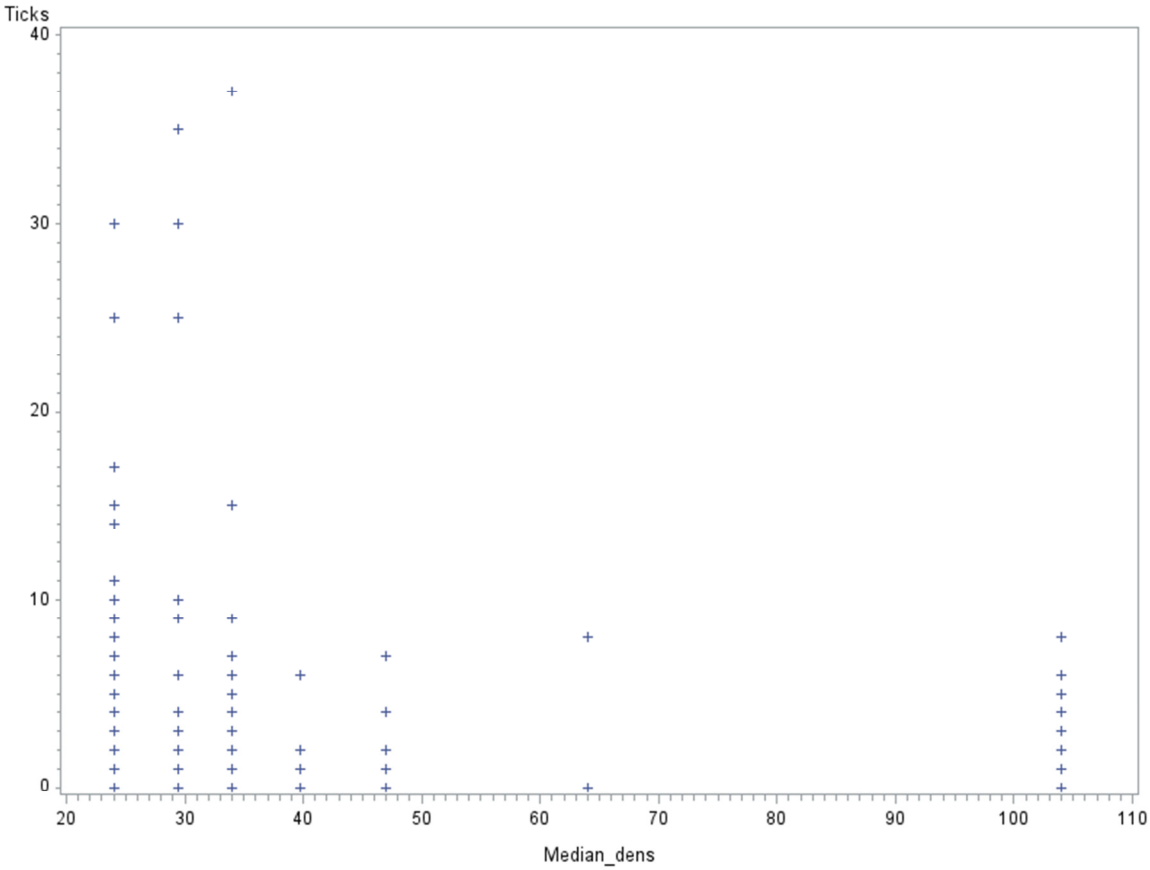


Figure 4.3. Tick counts on the right ear of hunted deer according to median deer density of park, 2001-2011

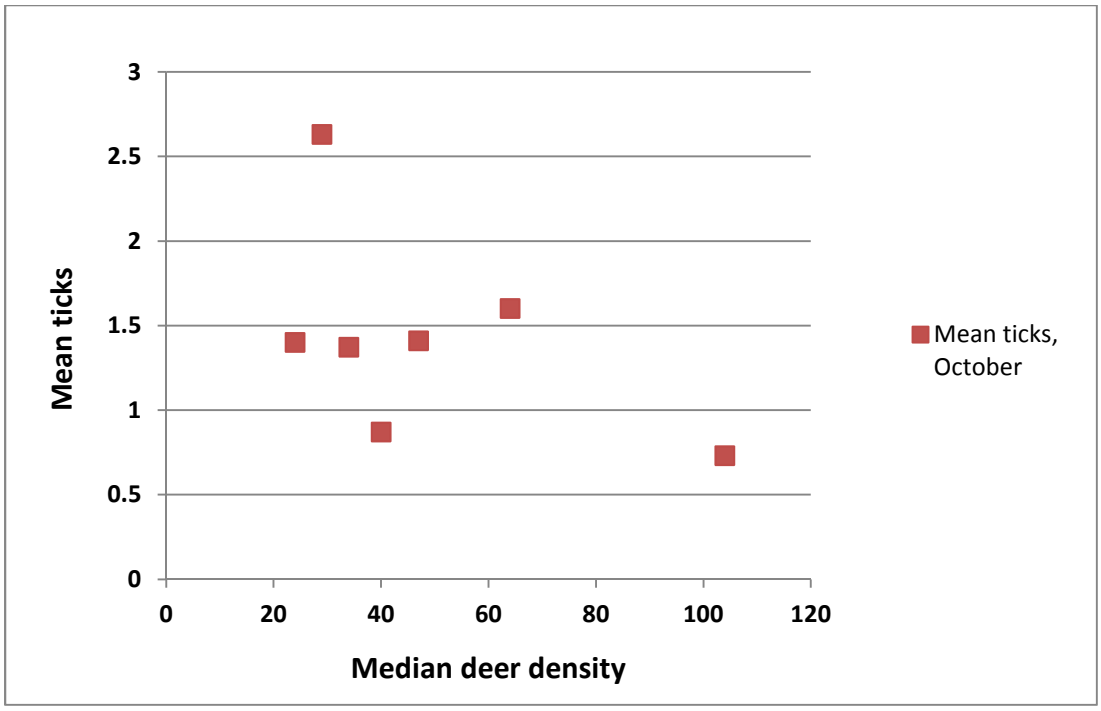


Figure 4.4. Mean number of ticks per deer and median deer density by park, October only

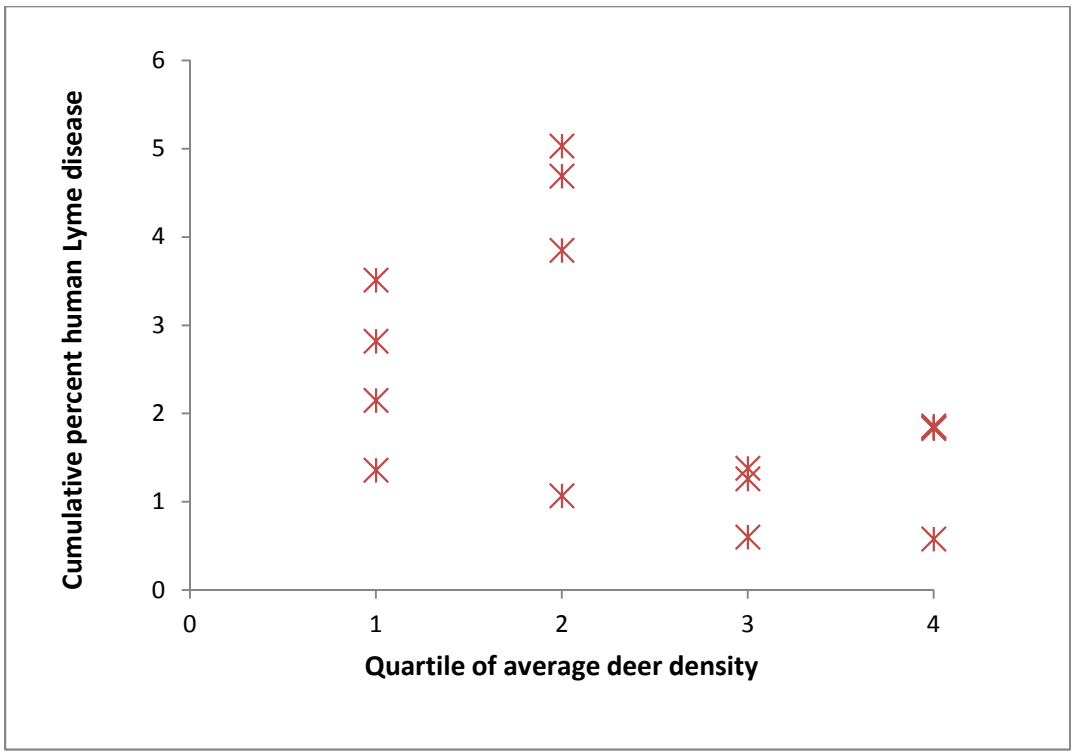


Figure 4.5. Cumulative incidence of human Lyme disease surrounding each park during 2001-2011, by quartile of deer density

CHAPTER 5: DISSERTATION DISCUSSION AND CONCLUSIONS

The number of human Lyme disease cases in the United States has continued to increase despite introduction of several possible prevention methods in the decades since the disease was first described (Stafford 2007; Centers for Disease Control and Prevention 2013a). In the absence of a commercially-available vaccine, environmental and human behavioral interventions are the only mechanisms by which to combat this tickborne disease. There is clear need for better prevention options, or for existing options to be more broadly and appropriately implemented, if the disease is to be controlled in the United States. Understanding factors that drive spatial variation in risk is helpful to appropriately implement or promote possible interventions.

The studies described in this dissertation quantified the spatial distribution of human Lyme disease in one county and demonstrated that a portion of spatial variation in disease risk was explained by environmental and socio-demographic factors. Project 1 demonstrated significant spatial clustering of human Lyme disease according to residence. Several environmental and socio-demographic differences were apparent between high-risk clusters and the remainder of Howard County; proportionally, high-risk clusters contained more low-density development and red and white oak forest, and more area characterized by higher average income and home value. Project 2 approached the spatial variation in risk of human Lyme disease in Howard County differently. Using multilevel models, significant variation in disease risk was evident between census block groups. Using different analysis methods and

observational units, many of the same factors were associated with elevated Lyme disease risk in Project 2 as were linked to high-risk clusters in Project 1 (i.e., residence in low- and medium-density development, and in area dominated by red and white oak forest and characterized by higher home values). Most of the variation in Lyme disease risk between census block groups was explained by these variables, although 25% of the variation in risk between census block groups remained unexplained. Most of the total variation in Lyme disease risk in Howard County occurred within (not between) census block groups, further illustrating the fine-scale nature of spatial variation in human Lyme disease risk. Nevertheless, Lyme disease risk in Howard County was clearly affected by factors acting at multiple levels.

Project 3 examined additional factors that might contribute to spatial variation in Lyme disease risk on a sub-county level, specifically variation in deer density and tick infestation on deer. These factors were examined separately and in a restricted subset of Howard County population because deer and tick data were not available across the entire county. In multivariable models, lower-density residential development and red and white oak forests were associated with disease risk after the effects of deer density and tick infestation on deer were controlled. The finding of higher tick infestation on deer in areas with lower deer density suggests that effects of moderate differences in deer density may be buffered— more ticks may attach to remaining deer in areas where deer are less plentiful. Yet, there were no areas in our study where deer were scarce. The role of deer density and tick infestation on deer in human disease risk is not well-understood. There are several other important factors which contribute to disease risk including abundance and infection prevalence of reservoir hosts, nymphal tick abundance, and human behavior, all of which were unmeasured in this analysis. This study

provided evidence that fine-scale differences in deer density are not clearly associated with spatial variation in Lyme disease risk within a single county; however, there were limitations to the completeness and accuracy of these data. Nevertheless, the study findings suggest that broad, community-driven moderate deer population reduction in inland settings may not be an effective Lyme disease prevention measure.

The studies described here were based on existing data. While more convenient, cheaper, and faster to obtain than prospectively designed and collected information, use of existing data leaves the investigator with several limitations that cannot be overcome including: type of data collected (e.g., tick counts on the right ear of deer), accuracy of the data collected (e.g., Lyme disease cases reported to the public health system), and artificial neighborhood boundaries for multilevel analysis (e.g., census block groups). The boundaries of census block groups used in these projects are likely not the most appropriate delineations for neighborhoods as they relate to risk of Lyme disease. Furthermore, true location of patient exposure is unknown for most cases. Yet, the demonstrated clustering by residential location suggests that residential address may be an appropriate proxy for most patients. However, whether the home itself, the immediate vicinity, or the neighborhood is where people in these areas most often contact infected ticks remains unresolved.

In epidemiologic studies, there is often a balance between validity and generalizability (Rothman et al. 2008). These studies were conducted using detailed environmental and Lyme disease patient address data in one single county. The level of detail in observation for both exposure and outcome would likely be sacrificed if one were to conduct this type of analysis

across a broader geographic area. As a consequence, although the estimates of effect in these studies are quite precise and relatively unbiased, there is limited ability to generalize these estimates or other findings to other geographic areas.

Many factors tied to human risk were constant across these studies. This consistency suggests these detected associations were not chance findings. Similar associations have been documented in other studies, using other designs, populations, and methods. Confirming these findings is informative, yet most of the detected associations do not necessarily provide clear, actionable information that can help prevent Lyme disease.

A decade ago Lubelczyk and colleagues (2004) suggested that refining the vegetation species associated with elevated tick abundance could help natural resource managers target management strategies (Lubelczyk et al. 2004). Exotic invasive shrub species have been tied to elevated tick abundance (Elias et al. 2006); the mechanism through which red and white oak forest may act on elevated disease risk could be through the underlying soil composition itself and leaf litter, or through increased presence of exotic invasive species in this type of forest. Further work should use a similar data set to assess the potential association between red and white oak forest and human disease in other places. This was the most common oak-dominated forest in this area, and it is possible that species specificity is meaningless.

These studies identified a potential association between forest structure and human Lyme disease risk. However, the proportion of the census block group composed of forest displayed the strongest association with Lyme disease risk in a multivariable setting that included residential land use and land cover characteristics. Conflicting data in the literature

suggest that blacklegged tick abundance and infection may be highest in small forest patches with more potential for human contact with edge environments and for reduced biodiversity to amplify; however, an opposite association with human disease has also been documented (higher incidence associated with larger forest patches)(Brownstein et al. 2005; Diuk-Wasser et al. 2012). Here, higher human risk associated with residence near larger average forest patches was demonstrated, but the association diminished in multivariable models. Understanding the mechanism through which forest structure and fragmentation affects tick abundance and infection prevalence is important to understanding Lyme disease ecology, but may be not be as important to disease prevention from a public health perspective. Humans likely encounter all types and shapes of forested environments in endemic areas, all of which have some infected ticks. If areas with higher entomologic risk are not where humans most frequently interact with forested environments, demonstration of higher tick abundance and infection prevalence in specific forest structures are irrelevant with regard to human Lyme disease.

While lower-density residential development was consistently associated with human risk in these projects, its role and relevance to prevention is unclear. Although low-density and sprawling development contributes to zoonotic disease emergence (Patz et al. 2004), modification of existing low-density development is probably not feasible. The mechanism through which low-density development increases disease risk is unknown, but it could reflect both environmental and behavioral factors. How individual and household landscape characteristics are tied to the environmental and socio-demographic factors we assessed is unknown, but critical to understand in order to appropriately implement prevention methods. Individual and household landscape-specific factors were unmeasured in these studies, and

likely play the largest role in placing humans at risk of Lyme disease in endemic areas. These factors include where people travel on a day to day basis, what types of forested habitats they may contact, and the nature of that contact. As further evidence of the importance of behavior, most of the spatial variation in risk in Lyme disease demonstrated in Project 2 was not between census block groups, but within. This finding underscores the very fine-scale variation in Lyme disease risk within endemic areas.

Future directions

In the decades since Lyme disease was first described, substantial ecologic research has contributed to understanding of the enzootic cycle. Yet, comparatively little research has addressed human risk of disease. The research community should redirect attention to understanding human behavior as it relates to contact with forested environments and Lyme disease risk. Furthermore, prevention methods need to be assessed according to whether human contact with ticks is actually reduced. Ultimately, although knowledge of the exact biological mechanisms through which disease occurs is valuable to design of control measures, disease prevention does not often require complete understanding of mechanism. Conversely, complete understanding of disease processes does not necessarily translate into effective prevention and control (Renwick 1973).

significant advances in computing power and statistical software in recent years have allowed for broader use of multilevel modeling methods, methodologic questions remain that relate to spatial relationships and logistic models. For example, logistic regression is common in health research, yet in a multilevel structure, measures of neighborhood importance are less

clear than in linear models—partitioning of variance is not possible as it is in linear models (Merlo et al. 2006). Furthermore, the intraclass correlation coefficient for multilevel logistic models is dependent upon the prevalence of the outcome, a feature which limits its interpretability and widespread use (Merlo et al. 2006). The median odds ratio provides a useful alternative tool to the intraclass correlation coefficient, but is not a widely understood measure, even among epidemiologists (Merlo et al. 2006). Finally, much multilevel research has ignored the spatial relationships among observations and groups of observations (Chaix et al. 2005a). Bridging of these disciplines is in its infancy, and appropriate implementation and interpretation remains a challenge. Future work in the immediate future should solidify the meaning and interpretability of neighborhood level variance, neighborhood level fixed-effect interpretation, and importance of residual spatial autocorrelation in order to properly interpret the importance of various levels in Lyme disease risk.

Although community-level interventions hold the most promise to prevent human Lyme disease because they lack reliance upon individual adherence, community deer management in inland settings may not be a viable and effective option; other strategies are drastically needed. The public health and scientific research community needs to identify viable and effective options for environmental intervention that can be accomplished at a broader scale in addition to efforts of individual homeowners. Although possibly managed by a county or town natural resource department, these theoretical community interventions would not be implemented evenly across their jurisdiction, but would be targeted to both public and private areas of the jurisdiction with highest risk. Multidisciplinary studies that incorporate individual behavior, landscape characteristics, reservoir abundance and infection prevalence, entomologic indices,

and human outcomes are critical, but scarce (Eisen and Eisen 2008; Finch et al. 2014). Advances in novel spatial epidemiologic analytic methods are needed to understand how the diverse, dynamic Lyme disease enzootic cycle and human behavior patterns are linked to Lyme disease risk across endemic regions before broad community-level interventions can become commonplace.

Conclusions

This dissertation advances knowledge of the fine-scale epidemiology of human Lyme disease and demonstrates the importance of using human outcome data in addition to entomologic data to understand variation in Lyme disease risk. This first known effort to quantify the sub-county spatial variation of human Lyme disease demonstrated that disease risk in some areas was twice that of other areas within the same county. Second, these studies, using advanced analytic methods, validated associations of previously recognized risk factors for Lyme disease—increased disease risk associated with low-density development and presence of forest. Moreover, this dissertation identified several associations with increased human Lyme disease risk that deserve further consideration—residence in areas composed of red and white oak forest, and that amount of forest in the vicinity of residence may be more important to disease risk than the fragmented structure of the forest. Additionally, Project 3 provided evidence that fine-scale variation in deer density in inland areas is not clearly and strongly associated with human risk. Consequently, community deer management programs, although helpful for deer overpopulation, may not be a viable Lyme disease prevention measure for inland areas. Finally, results of this dissertation suggest that multilevel models may help to provide insight regarding many remaining questions in the epidemiology of Lyme

disease; specifically these models may help define the scale at which various factors act on disease risk, and in turn, inform the most appropriate prevention methods.

CHAPTER 6: REFERENCES

Advisory Committee on Immunization Practices (1999). "Recommendations for the use of Lyme disease vaccine." Morbid. Mortal. Wkly. Rep. **48**(RR-7): 1-17.

Afzelius, A (1910). Verhandlungen der Dermatologischen Gesellschaft zu Stockholm. Arch Dermatol Syph.

Aguero-Rosenfeld, M, Wang, G, Schwartz, I and Wormser, G (2005). "Diagnosis of Lyme Borreliosis." Clin Microbiol Rev **18**(3): 484-509.

Akaike, H (1974). "New look at statistical-model identification." IEEE Trans Automat Contr Ac **19**: 716-723.

Alexandrescu, R, Jen, MH, Bottle, A, Jarman, B, et al. (2011). "Logistic versus hierarchical modeling: an analysis of a statewide inpatient sample." J Am Coll Surg **213**(3): 392-401.

Allan, BF, Keesing, F and Ostfeld, RS (2003). "Effect of forest fragmentation on Lyme disease risk." Conservation Biology **17**(1): 267-272.

Amerasinghe, FP, Breisch, NL, Azad, AF, Gimpel, WF, et al. (1992). "Distribution, density, and Lyme disease spirochete infection in *Ixodes dammini* (Acari: Ixodidae) on white-tailed deer in Maryland." J Med Entomol **29**(1): 54-61.

Amerasinghe, FP, Breisch, NL, Neidhardt, K, Pagac, B, et al. (1993). "Increasing density and *Borrelia burgdorferi* infection of deer-infesting *Ixodes dammini* (Acari: Ixodidae) in Maryland." J Med Entomol **30**(5): 858-864.

Anderson, JF and Magnarelli, LA (1980). "Vertebrate host relationships and distribution of ixodid ticks (Acari: Ixodidae) in Connecticut, USA." J Med Entomol **17**(4): 314-323.

Anderson, JF and Magnarelli, LA (1984). "Avian and mammalian hosts for spirochete-infected ticks and insects in a Lyme disease focus in Connecticut." Yale J Biol Med **57**(4): 627-641.

Apperson, CS, Levine, JF, Evans, TL, Braswell, A, et al. (1993). "Relative utilization of reptiles and rodents as hosts by immature *Ixodes scapularis* (Acari: Ixodidae) in the coastal plain of North Carolina, USA." Exp Appl Acarol **17**(10): 719-731.

Bacon, RM, Biggerstaff, BJ, Schriefer, ME, Gilmore, RD, Jr., et al. (2003). "Serodiagnosis of Lyme disease by kinetic enzyme-linked immunosorbent assay using recombinant VlsE1 or peptide antigens of *Borrelia burgdorferi* compared with 2-tiered testing using whole-cell lysates." J Infect Dis **187**(8): 1187-1199.

Bacon, RM, Kugeler, KJ and Mead, PS (2008). "Surveillance for Lyme disease--United States, 1992-2006." MMWR Surveill Summ **57**(10): 1-9.

Barbour, A (2012). "Remains of infection." J Clin Invest **122**(7): 2344-2346.

Barbour, AG (1998). "Fall and rise of Lyme disease and other *Ixodes* tick-borne infections in North America and Europe." Br Med Bull **54**(3): 647-658.

Bergstrom, S, Noppa, L, Gylfe, A and Ostberg, Y (2002). Molecular and Cellular Biology of *Borrelia burgdorferi* sensu lato. Lyme Borreliosis: Biology, Epidemiology, and Control. J Gray, O Kahl, RS Lane and G Stanek. New York, CABI**:** 47-90.

Bertille, N, Fournier-Charriere, E, Pons, G and Chalumeau, M (2013). "Managing fever in children: a national survey of parents' knowledge and practices in France." PLoS One **8**(12): e83469.

Brei, B, Brownstein, JS, George, JE, Pound, JM, et al. (2009). "Evaluation of the United States Department Of Agriculture Northeast Area-wide Tick Control Project by meta-analysis." Vector Borne Zoonotic Dis **9**(4): 423-430.

Brisson, D, Dykhuizen, DE and Ostfeld, RS (2008). "Conspicuous impacts of inconspicuous hosts on the Lyme disease epidemic." Proc Biol Sci **275**(1631): 227-235.

Brownstein, JS, Skelly, DK, Holford, TR and Fish, D (2005). "Forest fragmentation predicts local scale heterogeneity of Lyme disease risk." Oecologia **146**(3): 469-475.

Bunnell, JE, Price, SD, Das, A, Shields, TM, et al. (2003). "Geographic information systems and spatial analysis of adult *Ixodes scapularis* (Acari: Ixodidae) in the Middle Atlantic region of the U.S.A." J Med Entomol **40**(4): 570-576.

Burgdorfer, W, Barbour, AG, Hayes, SF, Benach, JL, et al. (1982). "Lyme disease-a tick-borne spirochetosis?" Science **216**(4552): 1317-1319.

Carey, AB, Krinsky, WL and Main, AJ (1980). "*Ixodes dammini* (Acari: Ixodidae) and associated ixodid ticks in south-central Connecticut, USA." J Med Entomol **17**(1): 89-99.

Carey, MG, Carey, AB, Main, AJ, Krinsky, WL, et al. (1981). "*Ixodes dammini* (Acari: Ixodidae) in forests in Connecticut." J Med Entomol **18**(2): 175-176.

Casjens, S, Palmer, N, van Vugt, R, WM, H, et al. (2000). "A bacterial genome in flux: the twelve linear and nine circular extrachromosomal DNAs in an infectious isolate of the Lyme disease spirochete *Borrelia burgdorferi*." Mol Microbiol **35**(3): 490-516.

Centers for Disease Control and Prevention (1995). "Recommendations for test performance and interpretation from the Second National Conference on Serologic Diagnosis of Lyme Disease." Morbid. Mortal. Wkly. Rep. **44**: 590-591.

Centers for Disease Control and Prevention. (2012). "National Notifiable Diseases Surveillance System (NNDSS)." Retrieved October 22, 2013, from http://wwwn.cdc.gov/nndss/default.aspx.

Centers for Disease Control and Prevention (2013a). "Summary of Notifiable Diseases--United States, 2011." Morbid. Mortal. Wkly. Rep. **60**(53).

Centers for Disease Control and Prevention (2013b). "Three sudden cardiac deaths associated with Lyme carditis --United States, November 2012-July 2013." Morbid. Mortal. Wkly. Rep. **62**(49): 993-996.

Chaix, B, Merlo, J and Chauvin, P (2005a). "Comparison of a spatial approach with the multilevel approach for investigating place effects on health: the example of healthcare utilisation in France." J Epidemiol Community Health **59**(6): 517-526.

Chaix, B, Merlo, J, Subramanian, SV, Lynch, J, et al. (2005b). "Comparison of a spatial perspective with the multilevel analytical approach in neighborhood studies: the case of mental and behavioral disorders due to psychoactive substance use in Malmo, Sweden, 2001." Am J Epidemiol **162**(2): 171-182.

Chandra, A, Latov, N, Wormser, GP, Marques, AR, et al. (2011a). "Epitope mapping of antibodies to VlsE protein of *Borrelia burgdorferi* in post-Lyme disease syndrome." Clin Immunol **141**(1): 103-110.

Chandra, A, Wormser, GP, Klempner, MS, Trevino, RP, et al. (2010). "Anti-neural antibody reactivity in patients with a history of Lyme borreliosis and persistent symptoms." Brain Behav Immun **24**(6): 1018-1024.

Chandra, A, Wormser, GP, Marques, AR, Latov, N, et al. (2011b). "Anti-*Borrelia burgdorferi* antibody profile in post-Lyme disease syndrome." Clin Vaccine Immunol **18**(5): 767-771.

Chen, H, Stratton, HH, Caraco, TB and White, DJ (2006). "Spatiotemporal Bayesian analysis of Lyme disease in New York State, 1990-2000." J Med Entomol **43**(4): 777-784.

Chen, H, White, DJ, Caraco, TB and Stratton, HH (2005). "Epidemic and spatial dynamics of Lyme disease in New York State, 1990-2000." J Med Entomol **42**(5): 899-908.

Connally, NP, Durante, AJ, Yousey-Hindes, KM, Meek, JI, et al. (2009). "Peridomestic Lyme disease prevention: results of a population-based case-control study." Am J Prev Med **37**(3): 201-206.

Connally, NP, Ginsberg, HS and Mather, TN (2006). "Assessing peridomestic entomological factors as predictors for Lyme disease." J Vector Ecol **31**(2): 364-370.

Conover, M (2002). Resolving Human-Wildlife Conflicts: the Science of Wildlife Damage Management. Boca Raton, Lewis Publishers.

Cortinas, MR and Kitron, U (2006). "County-level surveillance of white-tailed deer infestation by *Ixodes scapularis* and *Dermacentor albipictus* (Acari: Ixodidae) along the Illinois River." J Med Entomol **43**(5): 810-819.

Council of State and Territorial Epidemiologists. (1996). "Lyme disease case definition." Retrieved March 5, 2014, from http://wwwn.cdc.gov/nndss/script/conditionsummary.aspx?CondID=100.

Council of State and Territorial Epidemiologists. (2008). "Lyme disease case definition." Retrieved March 5, 2014, from http://wwwn.cdc.gov/nndss/script/conditionsummary.aspx?CondID=100.

Cromley, EK, Cartter, M, Mrozinski, RD and Ertel, S (1998). "Residential setting as a risk factor for Lyme disease in a hyperendemic region." American Journal of Epidemiology **147**(5): 472-477.

Daniels, TJ and Fish, D (1995). "Effect of deer exclusion on the abundance of immature *Ixodes scapularis* (Acari: Ixodidae) parasitizing small and medium-sized mammals." J Med Entomol **32**(1): 5-11.

Daniels, TJ, Fish, D, Levine, JF, Greco, MA, et al. (1993a). "Canine exposure to *Borrelia burgdorferi* and prevalence of *Ixodes dammini* (Acari: Ixodidae) on deer as a measure of Lyme disease risk in the northeastern United States." J Med Entomol **30**(1): 171-178.

Daniels, TJ, Fish, D and Schwartz, I (1993b). "Reduced abundance of *Ixodes scapularis* (Acari: Ixodidae) and Lyme disease risk by deer exclusion." J Med Entomol **30**(6): 1043-1049.

Deblinger, RD, Wilson, ML, Rimmer, DW and Spielman, A (1993). "Reduced abundance of immature *Ixodes dammini* (Acari: Ixodidae) following incremental removal of deer." J Med Entomol **30**(1): 144-150.

Diez-Roux, AV (1998). "Bringing context back into epidemiology: variables and fallacies in multilevel analysis." Am J Public Health **88**(2): 216-222.

Diez-Roux, AV (2000). "Multilevel analysis in public health research." Annu Rev Public Health **21**: 171-192.

Diez Roux, AV (2002). "A glossary for multilevel analysis." J Epidemiol Community Health **56**(8): 588-594.

Diez Roux, AV and Aiello, AE (2005). "Multilevel analysis of infectious diseases." J Infect Dis **191 Suppl 1**: S25-33.

Dister, SW, Fish, D, Bros, SM, Frank, DH, et al. (1997). "Landscape characterization of peridomestic risk for Lyme disease using satellite imagery." Am J Trop Med Hyg **57**(6): 687-692.

Diuk-Wasser, MA, Hoen, AG, Cislo, P, Brinkerhoff, R, et al. (2012). "Human risk of infection with *Borrelia burgdorferi*, the Lyme disease agent, in eastern United States." Am J Trop Med Hyg **86**(2): 320-327.

Donahue, JG, Piesman, J and Spielman, A (1987). "Reservoir competence of white-footed mice for Lyme disease spirochetes." Am. J. Trop. Med. Hyg. **36**: 92-96.

Doyle, TJ, Glynn, MK and Groseclose, LS (2002). "Completeness of notifiable infectious disease reporting in the United States." Am J Epidemiol **155**: 866-874.

Duffy, DC, Campbell, SR, Clark, D, DiMotta, C, et al. (1994). "*Ixodes scapularis* (Acari: Ixodidae) deer tick mesoscale populations in natural areas: effects of deer, area, and location." J Med Entomol **31**(1): 152-158.

Duray, PH and Steere, AC (1988). "Clinical pathologic correlations of Lyme disease by stage." Ann N Y Acad Sci **539**: 65-79.

Eisen, L and Eisen, RJ (2007). "Need for improved methods to collect and present spatial epidemiologic data for vector-borne diseases." Emerg Infect Dis **13**(12): 1816-1820.

Eisen, L and Eisen, RJ (2011). "Using geographic information systems and decision support systems for the prediction, prevention, and control of vector-borne diseases." Annu Rev Entomol **56**: 41-61.

Eisen, RJ and Eisen, L (2008). "Spatial modeling of human risk of exposure to vector-borne pathogens based on epidemiological versus arthropod vector data." J Med Entomol **45**(2): 181-192.

Eisen, RJ, Lane, RS, Fritz, CL and Eisen, L (2006). "Spatial patterns of Lyme disease risk in California based on disease incidence data and modeling of vector-tick exposure." Am J Trop Med Hyg **75**(4): 669-676.

Eisen, RJ, Piesman, J, Zielinski-Gutierrez, E and Eisen, L (2012). "What do we need to know about disease ecology to prevent Lyme disease in the northeastern United States?" J Med Entomol **49**(1): 11-22.

Elias, SP, Lubelczyk, CB, Rand, PW, Lacombe, EH, et al. (2006). "Deer browse resistant exotic-invasive understory: an indicator of elevated human risk of exposure to *Ixodes scapularis* (Acari: Ixodidae) in southern coastal Maine woodlands." J Med Entomol **43**(6): 1142-1152.

Falco, RC and Fish, D (1988). "Prevalence of *Ixodes dammini* near the homes of Lyme disease patients in Westchester County, New York." Am J Epidemiol **127**: 826-830.

Falco, RC, Fish, D and Piesman, J (1996). "Duration of tick bites in a Lyme disease-endemic area." Am J Epidemiol **143**(2): 187-192.

Falco, RC, McKenna, DF, Daniels, TJ, Nadelman, RB, et al. (1999). "Temporal relation between *Ixodes scapularis* abundance and risk for Lyme disease associated with erythema migrans." Am J Epidemiol **149**(8): 771-776.

Feder, HM, Jr., Gerber, MA, Luger, SW and Ryan, RW (1992). "Persistence of serum antibodies to *Borrelia burgdorferi* in patients treated for Lyme disease." Clin Infect Dis **15**(5): 788-793.

Finch, C, Al-Damluji, MS, Krause, PJ, Niccolai, L, et al. (2014). "Integrated assessment of behavioral and environmental risk factors for Lyme disease infection on Block Island, Rhode Island." PLoS One **9**(1): e84758.

Fish, D and Childs, JE (2009). "Community-based prevention of Lyme disease and other tick-borne diseases through topical application of acaricide to white-tailed deer: background and rationale." Vector Borne Zoonotic Dis **9**(4): 357-364.

Frank, C, Fix, AD, Pena, CA and Strickland, GT (2002). "Mapping Lyme disease incidence for diagnostic and preventive decisions, Maryland." Emerg Infect Dis **8**(4): 427-429.

Fraser, CM, Casjens, S, Huang, WM, Sutton, GG, et al. (1997). "Genomic sequence of a Lyme disease spirochaete, *Borrelia burgdorferi*." Nature **390**(6660): 580-586.

French, JB, Jr., Schell, WL, Kazmierczak, JJ and Davis, JP (1992). "Changes in population density and distribution of *Ixodes dammini* (Acari: Ixodidae) in Wisconsin during the 1980s." J Med Entomol **29**(5): 723-728.

Gardner, W, Mulvey, EP and Shaw, EC (1995). "Regression analyses of counts and rates: Poisson, overdispersed Poisson, and negative binomial models." Psychol Bull **118**(3): 392-404.

Garnett, JM, Connally, NP, Stafford, KC, 3rd and Cartter, ML (2011). "Evaluation of deer-targeted interventions on Lyme disease incidence in Connecticut." Public Health Rep **126**(3): 446-454.

Ginsberg, HS and Zhioua, E (1999). "Influence of deer abundance on the abundance of questing adult *Ixodes scapularis* (Acari: Ixodidae)." J Med Entomol **36**(3): 376-381.

Ginsberg, HS, Zhioua, E, Mitra, S, Fischer, J, et al. (2004). "Woodland type and spatial distribution of nymphal *Ixodes scapularis* (Acari: Ixodidae) " Environmental Entomology **33**(5): 1266-1273.

Glass, GE, Amerasinghe, FP, Morgan, JM, 3rd and Scott, TW (1994). "Predicting *Ixodes scapularis* abundance on white-tailed deer using geographic information systems." Am J Trop Med Hyg **51**(5): 538-544.

Glass, GE, Schwartz, BS, Morgan, JM, Johnson, DT, et al. (1995). "Environmental risk factors for Lyme disease identified with geographic information systems." American Journal of Public Health **85**(7): 944-948.

Glavanakov, S, White, D, Caraco, T, Lapenis, A, et al. (2001). "Lyme disease in New York state: spatial pattern at a regional scale " Am J Trop Med Hyg **65**(5): 538-545.

Goddard, J (2003). Physician's Guide to Arthropods of Medical Importance, CRC Press.

Gould, LH, Nelson, RS, Griffith, KS, Hayes, EB, et al. (2008). "Knowledge, attitudes, and behaviors regarding Lyme disease prevention among Connecticut residents, 1999-2004." Vector Borne Zoonotic Dis **8**(6): 769-776.

Guerra, M, Walker, E, Jones, C, Paskewitz, S, et al. (2002). "Predicting the risk of Lyme disease: habitat suitability for *Ixodes scapularis* in the north central United States." Emerg Infect Dis **8**(3): 289-297.

Hanley, JA, Negassa, A, Edwardes, MD and Forrester, JE (2003). "Statistical analysis of correlated data using generalized estimating equations: an orientation." Am J Epidemiol **157**(4): 364-375.

Hayes, EB and Piesman, J (2003). "How can we prevent Lyme disease?" N Engl J Med **348**(24): 2424-2430.

Herrington, JE (2004). "Risk perceptions regarding ticks and Lyme disease: a national survey." Am J Prev Med **26**(2): 135-140.

Hinckley, AF, Connally, NP, Meek, JI, Johnson, BJ, et al. (2014). "Lyme disease testing by large commercial laboratories in the United States." Clin Infect Dis(doi: 10.1093/cid/ciu397).

Hoffman, JP (2004). Generalized Linear Models: An Applied Approach. Boston, Massachusetts, Pearson.

Hojgaard, A, Lukacik, G and Piesman, J (2014). "Detection of *Borrelia burgdorferi*, *Anaplasma phagocytophilum,* and *Babesia microti,* with two different multiplex PCR assays." Ticks Tick Borne Dis **5**(3): 349-351.

Hosmer, DW and Lemeshow, S (2000). Applied Logistic Regression. New York, John Wiley and Sons.

Howard County Department of Recreation and Parks. (2000). "Comprehensive Deer Management Plan." Retrieved April 1 2011, from http://countyofhowardmd.us/DisplayPrimary.aspx?id=6442456013.

Hubbard, AE, Ahern, J, Fleischer, NL, Van der Laan, M, et al. (2010). "To GEE or not to GEE: comparing population average and mixed models for estimating the associations between neighborhood risk factors and health." Epidemiology **21**(4): 467-474.

Jackson, LE, Hilborn, ED and Thomas, JC (2006a). "Towards landscape design guidelines for reducing Lyme disease risk." International Journal of Epidemiology **35**: 315-322.

Jackson, LE, Levine, JF and Hilborn, ED (2006b). "A comparison of analysis units for associating Lyme disease with forest-edge habitat." Community Ecology **7**(2): 189-197.

Jones, CG, Ostfeld, RS, Richard, MP, Schauber, EM, et al. (1998a). "Chain reactions linking acorns to gypsy moth outbreaks and Lyme disease risk." Science **279**(5353): 1023-1026.

Jones, CG, Ostfeld, RS, Richard, MP, Schauber, EM, et al. (1998b). "Mast seeding and Lyme disease." Trends Ecol Evol **13**(12): 506.

Jordan, RA and Schulze, TL (2005). "Deer browsing and the distribution of *Ixodes scapularis* (Acari: Ixodidae) in Central New Jersey forests." Environmental Entomology **34**(4): 801-806.

Jordan, RA, Schulze, TL and Jahn, MB (2007). "Effects of reduced deer density on the abundance of *Ixodes scapularis* (Acari: Ixodidae) and Lyme disease incidence in a northern New Jersey endemic area." J Med Entomol **44**(5): 752-757.

Kahl, O, Gern, L, Eisen, L and Lane, RS (2002). Ecological research on *Borrelia burgdorferi* sensu lato: terminology and some methodological pitfalls. Lyme Borreliosis: Biology, Epidemiology and Control. JS Gray, O Kahl, RS Lane and G Stanek. New York, CABI.

Keefe, LM, Moro, MH, Vinasco, J, Hill, C, et al. (2009). "The use of harvested white-tailed deer *(Odocoileus virginianus)* and geographic information system (GIS) methods to characterize distribution and locate spatial clusters of *Borrelia burgdorferi* and its vector *Ixodes scapularis* in Indiana." Vector Borne Zoonotic Dis **9**(6): 671-680.

Keesing, F, Brunner, J, Duerr, S, Killilea, M, et al. (2009). "Hosts as ecological traps for the vector of Lyme disease." Proc Biol Sci **276**(1675): 3911-3919.

Killilea, ME, Swei, A, Lane, RS, Briggs, CJ, et al. (2008). "Spatial dynamics of Lyme disease: a review." Ecohealth **5**(2): 167-195.

Kilpatrick, HJ and LaBonte, AM (2003). "Deer hunting in a residential community: the community's perspective." Wildlife Society Bulletin **31**(340-348).

Kilpatrick, HJ and Spohr, SM (2000). "Spatial and temporal use of a suburban landscape by female white-tailed deer." Wildlife Society Bulletin **28**(4): 1023-1029.

Kilpatrick, HJ and Walter, WD (1997). "Urban Deer Management: A Community Vote." Wildlife Society Bulletin **25**(2): 388-391.

Kitron, U (1998). "Landscape ecology and epidemiology of vector-borne diseases: tools for spatial analysis." J Med Entomol **35**(4): 435-445.

Kitron, U, Jones, CJ, Bouseman, JK, Nelson, JA, et al. (1992). "Spatial analysis of the distribution of *Ixodes dammini* (Acari: Ixodidae) on white-tailed deer in Ogle County, Illinois." J Med Entomol **29**(2): 259-266.

Kitron, U and Kazmierczak, J (1997). "Spatial analysis of the distribution of Lyme disease in Wisconsin." American Journal of Epidemiology **145**(6): 558-565.

Klein, JD, Eppes, SC and Hunt, P (1996). "Environmental and life-style risk factors for Lyme disease in children." Clin Pediatr (Phila) **35**(7): 359-363.

Kugeler, KJ, Griffith, KS, Gould, LH, Kochanek, K, et al. (2011). "A review of death certificates listing Lyme disease as a cause of death in the United States." Clin Infect Dis **52**(3): 364-367.

Kulldorff M. and Information Management Services Inc. SaTScan: software for the spatial and space-time scan statistics. www.satscan.org, 2011.

Kurtenbach, K, Sewell, HS, Ogden, NH, Randolph, SE, et al. (1998). "Serum complement sensitivity as a key factor in Lyme disease ecology." Infect Immun **66**(3): 1248-1251.

Lacombe, E, Rand, PW and Smith, RP, Jr. (1993). "Disparity of *Borrelia burgdorferi* infection rates of adult *Ixodes dammini* on deer and vegetation." J Infect Dis **167**(5): 1236-1238.

Lastavica, CC, Wilson, ML, Berardi, VP, Spielman, A, et al. (1989). "Rapid emergence of a focal epidemic of Lyme disease in coastal Massachusetts." New Engl J Med **320**(3): 133-137.

Levi, T, Kilpatrick, AM, Mangel, M and Wilmers, CC (2012). "Deer, predators, and the emergence of Lyme disease." Proc Natl Acad Sci U S A **109**(27): 10942-10947.

Levine, JF, Wilson, ML and Spielman, A (1985). "Mice as reservoirs of the Lyme disease spirochete." Am J Trop Med Hyg **34**(355-360).

Ley, C, Olshen, EM and Reingold, AL (1995). "Case-control study of risk factors for incident Lyme disease in California." American Journal of Epidemiology **142**(9 Suppl): S39-47.

LoGiudice, K, Duerr, ST, Newhouse, MJ, Schmidt, KA, et al. (2008). "Impact of host community composition on Lyme disease risk." Ecology **89**(10): 2841-2849.

LoGiudice, K, Ostfeld, RS, Schmidt, KA and Keesing, F (2003). "The ecology of infectious disease: effects of host diversity and community composition on Lyme disease risk." Proc Natl Acad Sci U S A **100**(2): 567-571.

Lubelczyk, CB, Elias, SP, Rand, PW, Holman, MS, et al. (2004). "Habitat associations of *Ixodes scapularis* (Acari: Ixodidae) in Maine." Environmental Entomology **33**: 900-906.

Maas, C and Hox, J (2005). "Sufficient sample sizes for multilevel modeling." Methodology **1**: 86-92.

Main, AJ, Carey, AB, Carey, MG and Goodwin, RH (1982). "Immature *Ixodes dammini* (Acari: Ixodidae) on small animals in Connecticut, USA." J Med Entomol **19**(6): 655-664.

Main, AJ, Sprance, HE, Kloter, KO and Brown, SE (1981). "*Ixodes dammini* (Acari: Ixodidae) on white-tailed deer (*Odocoileus virginianus*) in Connecticut." Journal of Medical Entomology **18**(6): 487-492.

Maldonado, G and Greenland, S (1993). "Simulation study of confounder-selection strategies." Am J Epidemiol **138**(11): 923-936.

Mather, TN and Mather, ME (1990). "Intrinsic competence of three ixodid ticks (Acari) as vectors of the Lyme disease spirochete." J Med Entomol **27**(4): 646-650.

Mather, TN, Telford, SR, 3rd, Moore, SI and Spielman, A (1990). "*Borrelia burgdorferi* and *Babesia microti*: efficiency of transmission from reservoirs to vector ticks *(Ixodes dammini)*." Exp Parasitol **70**(1): 55-61.

Mather, TN, Wilson, ML, Moore, SI, Ribeiro, JM, et al. (1989). "Comparing the relative potential of rodents as reservoirs of the Lyme disease spirochete (*Borrelia burgdorferi*)." American Journal of Epidemiology **130**(1): 143-150.

Matuschka, FR, Ohlenbusch, A, Eiffert, H, Richter, D, et al. (1996). "Characteristics of Lyme disease spirochetes in archived European ticks." Journal of Infectious Diseases **174**: 424-426.

Maupin, GO, Fish, D, Zultowsky, J, Campos, EG, et al. (1991). "Landscape ecology of Lyme disease in a residential area of Westchester County, New York." Am J Epidemiol **133**(11): 1105-1113.

McEnroe, WD (1977). "The restriction of the species range of *Ixodes scapularis* (Say) in Massachusetts by fall and winter temperature." Acarologia **18**: 618-625.

Meek, JI, Roberts, CL, Smith, EV, Jr. and Cartter, ML (1996). "Underreporting of Lyme disease by Connecticut physicians, 1992." J Public Health Manag Pract **2**(4): 61-65.

Merlo, J, Chaix, B, Ohlsson, H, Beckman, A, et al. (2006). "A brief conceptual tutorial of multilevel analysis in social epidemiology: using measures of clustering in multilevel logistic regression to investigate contextual phenomena." J Epidemiol Community Health **60**(4): 290-297.

Merlo, J, Chaix, B, Yang, M, Lynch, J, et al. (2005). "A brief conceptual tutorial on multilevel analysis in social epidemiology: interpreting neighbourhood differences and the effect of neighbourhood characteristics on individual health." J Epidemiol Community Health **59**(12): 1022-1028.

Moineddin, R, Matheson, FI and Glazier, RH (2007). "A simulation study of sample size for multilevel logistic regression models." BMC Med Res Methodol **7**: 34.

Morgenstern, H (2008). Ecologic Studies. Modern Epidemiology. Philadephia, Lippincott, Williams & Wilkins**:** 511-532.

Nadelman, RB and Wormser, GP (1998). "Lyme borreliosis." Lancet **352**(9127): 557-565.

Naleway, AL, Belongia, EA, Kazmierczak, JJ, Greenlee, RT, et al. (2002). "Lyme disease incidence in Wisconsin: a comparison of state-reported rates and rates from a population-based cohort." Am J Epidemiol **155**(12): 1120-1127.

New Jersey Department of Environmental Protection. (2010). "Biology of the white-tailed deer *Odocoileus virginianus*." Retrieved September 18, 2012, from http://www.nj.gov/dep/fgw/deerbiol.htm.

Nicholson, MC and Mather, TN (1996). "Methods for evaluating Lyme disease risks using geographic information systems and geospatial analysis." Journal of Medical Entomology **33**(5): 711-720.

Northeast Deer Technical Committee. (2009). "An Evaluation of Deer Management Options." from http://www.nj.gov/dep/fgw/pdf/deer_mgt_options.pdf.

Oliver, JHJ, Cummins, GA and Joiner, MS (1993). "Immature *Ixodes scapularis* (Acari: Ixodidae) parasitizing lizards from the southeastern U.S.A." J Parasitol **79**(5): 684-689.

Orloski, KA, Campbell, GL, Genese, CA, Beckley, JW, et al. (1998). "Emergence of Lyme disease in Hunterdon County, New Jersey, 1993: a case-control study of risk factors and evaluation of reporting patterns." Am J Epidemiol **147**(4): 391-397.

Ostfeld, RS (2011). Lyme disease: the ecology of a complex system. New York, Oxford.

Ostfeld, RS, Canham, CD, Oggenfuss, K, Winchcombe, RJ, et al. (2006). "Climate, deer, rodents, and acorns as determinants of variation in Lyme disease risk." PLoS Biol **4**(6): e145.

Ostfeld, RS, Cepeda, OM, Hazler, KR and Miller, MC (1995). "Ecology of Lyme disease: habitat associations of ticks (*Ixodes scapularis*) in a rural landscape." Ecol Appl **5**(2): 353-361.

Ostfeld, RS, Schauber, EM, Canham, CD, Keesing, F, et al. (2001). "Effects of acorn production and mouse abundance on abundance and *Borrelia burgdorferi* infection prevalence of nymphal *Ixodes scapularis* ticks." Vector Borne Zoonotic Dis **1**(1): 55-63.

Ott, RL and Longnecker, M (2001). Introduction to Statistical Methods and Data Analysis. Pacific Grove, CA, Duxbury.

Palm, TA (1890). "The geographical distribution and aetiology of rickets." Practitioner **45**: 270-279.

Pan, W (2001). "Akaike's information criterion in generalized estimating equations." Biometrics **57**: 120-125.

Pardanani, N and Mather, TN (2004). "Lack of spatial autocorrelation in fine-scale distributions of *Ixodes scapularis* (Acari: Ixodidae)." J Med Entomol **41**(5): 861-864.

Patz, JA, Daszak, P, Tabor, GM, Aguirre, AA, et al. (2004). "Unhealthy landscapes: Policy recommendations on land use change and infectious disease emergence." Environ Health Perspect **112**(10): 1092-1098.

Persing, DH, Telford III, SR, Rys, PN, Dodge, DE, et al. (1990). "Detection of *Borrelia burgdorferi* DNA in museum specimens of *Ixodes dammini* ticks." Science **249**(4975): 1420-1423.

Peterson, I, Borrell, LN, El-Sadr, W and Teklehaimanot, A (2009). "A temporal-spatial analysis of malaria transmission in Adama, Ethiopia." Am J Trop Med Hyg **81**(6): 944-949.

Pichon, B, Mousson, L, Figureau, C, Rodhain, F, et al. (1999). "Density of deer in relation to the prevalence of *Borrelia burgdorferi* s.l. in *Ixodes ricinus* nymphs in Rambouillet forest, France." Exp Appl Acarol **23**(3): 267-275.

Piesman, J (2002). Ecology of *Borrelia burgdorferi* in North America. Lyme Borreliosis: Biology, Epidemiology, and Control. JS Gray, O Kahl, RS Lane and G Stanek. New York, CABI.

Piesman, J, Donahue, JG, Mather, TN and Spielman, A (1986). "Transovarially acquired Lyme disease spirochetes *(Borrelia burgdorferi)* in field-collected larval *Ixodes dammini* (Acari: Ixodidae)." J Med Entomol **23**(2): 219.

Piesman, J and Gern, L (2004). "Lyme Borreliosis in Europe and North America." Parasitology **129**: S191-220.

Piesman, J, Mather, TN, Dammin, GJ, Telford III, SR, et al. (1987). "Seasonal variation of transmission risk of Lyme disease and human babesiosis." Am J Epidemiol **126**: 1187-1189.

Piesman, J, Maupin, GO, Campos, EG and Happ, CM (1991). "Duration of adult female *Ixodes dammini* attachment and transmission of *Borrelia burgdorferi*, with description of a needle aspiration isolation method." J Infect Dis **163**(4): 895-897.

Piesman, J and Sinsky, RJ (1988). "Ability of *Ixodes scapularis*, *Dermacentor variabilis*, and *Amblyomma americanum* (Acari: Ixodidae) to acquire, maintain, and transmit Lyme disease spirochetes *(Borrelia burgdorferi)*." J Med Entomol **25**(5): 336-339.

Piesman, J and Spielman, A (1979). "Host-associations and seasonal abundance of immature *Ixodes dammini* in southeastern Massachusetts." Annals Entomological Soc. America **72**: 829-832.

Piesman, J, Spielman, A, Etkind, P, Ruebush, TK, 2nd, et al. (1979). "Role of deer in the epizootiology of *Babesia microti* in Massachusetts, USA." J Med Entomol **15**(5-6): 537-540.

Raizman, EA, Holland, JD and Shukle, JT (2012). "White-tailed deer *(Odocoileus virginianus)* as a potential sentinel for human Lyme disease in Indiana." Zoonoses Public Health **60**(3): 227-233.

Rand, PW, Lubelczyk, C, Holman, MS, Lacombe, EH, et al. (2004). "Abundance of *Ixodes scapularis* (Acari: Ixodidae) after the complete removal of deer from an isolated offshore island, endemic for Lyme disease." J Med Entomol **41**(4): 779-784.

Rand, PW, Lubelczyk, C, Lavigne, GR, Elias, S, et al. (2003). "Deer density and the abundance of *Ixodes scapularis* (Acari: Ixodidae)." J Med Entomol **40**(2): 179-184.

Rempel, RS, Kaukinen, D and Carr, AP Patch Analyst and Patch Grid. http://www.cnfer.on.ca/SEP/patchanalyst/index.html, 2012.

Renwick, JH (1973). "Analysis of cause--Long cut to prevention?" Nature **246**: 114-115.

Rothman, KJ, Greenland, S and Lash, TL (2008). Modern Epidemiology. Philadelphia, Lippincott Williams & Williams.

Schabenberger, O and Gotway, CA (2005). Statistical Methods in Spatial Data Analysis. Boca Raton, CRC Press.

Schulze, TL and Jordan, RA (1995). "Potential influence of leaf litter depth on effectiveness of granular carbaryl against subadult *Ixodes scapularis* (Acari: Ixodidae)." J Med Entomol **32**(2): 205-208.

Schulze, TL, Jordan, RA and Hung, RW (2001a). "Potential effects of animal activity on the spatial distribution of *Ixodes scapularis* and *Amblyomma americanum* (Acari: Ixodidae)." Environmental Entomology **30**(3): 568-577.

Schulze, TL, Jordan, RA, Hung, RW, Taylor, RC, et al. (2001b). "Efficacy of granular deltamethrin against *Ixodes scapularis* and *Amblyomma americanum* (Acari: Ixodidade) nymphs." J Med Entomol **38**(2): 344-346.

Schulze, TL, Jordan, RA and Krivenko, AJ (2005). "Effects of barrier application of granular deltamethrin on subadult *Ixodes scapularis* (Acari: Ixodidae) and nontarget forest floor arthropods." J Econ Entomol **98**(3): 976-981.

Schulze, TL, Jordan, RA, Schulze, CJ and Healy, SP (2008). "Suppression of *Ixodes scapularis* (Acari: Ixodidae) following annual habitat-targeted acaricide applications against fall populations of adults." J Am Mosq Control Assoc **24**(4): 566-570.

Schulze, TL, Lakat, MF, Bowen, GS, Parkin, WE, et al. (1984). "*Ixodes dammini* (Acari: Ixodidae) and other ixodid ticks collected from white-tailed deer in New Jersey, USA. I. Geographical distribution and its relation to selected environmental and physical factors." J Med Entomol **21**(6): 741-749.

Schulze, TL, Taylor, GC, Jordan, RA, Bosler, EM, et al. (1991). "Effectiveness of selected granular acaricide formulations in suppressing populations of *Ixodes dammini* (Acari: Ixodidae): short-term control of nymphs and larvae." J Med Entomol **28**(5): 624-629.

Scrimenti, RJ (1970). "Erythema chronicum migrans." Arch Dermatol **102**: 104-105.

Sikand, V and Muellegger, RR (2011). Early Lyme Borreliosis. Lyme Borreliosis in Europe and North America: Epidemiology and Clinical Practice. SK Sood. Hoboken New Jersey, John Wiley and Sons.

Singh, S, Lin, YL, Kuo, YF, Nattinger, AB, et al. (2013). "Variation in the risk of readmission among hospitals: the relative contribution of patient, hospital and inpatient provider characteristics." J Gen Intern Med.

Slajchert, T, Kitron, UD, Jones, CJ and Mannelli, A (1997). "Role of the eastern chipmunk *(Tamias striatus)* in the epizootiology of Lyme borreliosis in northwestern Illinois, USA." J Wildl Dis **33**(1): 40-46.

Smith, G, Wileyto, EP, Hopkins, RB, Cherry, BR, et al. (2001). "Risk factors for Lyme disease in Chester County, Pennsylvania." Public Health Rep **116 Suppl 1**: 146-156.

Smith, RP, Schoen, RT, Rahn, DW, Sikand, VK, et al. (2002). "Clinical characteristics and treatment outcome of early Lyme disease in patients with microbiologically confirmed erythema migrans." Ann Intern Med **136**(6): 421-428.

Snijders, TAB and Bosker, RJ (1999). Multilevel analysis: an introduction to basic and advanced multi-level modeling. Thousand Oaks, Sage Publishers.

Snow, J (1855). On the Mode of Communication of Cholera. London, John Churchill.

Sparrowe, RD and Springer, PF (1970). "Seasonal activity patterns of white-tailed deer in eastern South Dakota." Journal of Wildlife Management **34**: 420-431.

Stafford, KC (2007). Tick Management Handbook: An integrated guide for homeowners, pest control operators, and public health officials for the prevention of tick-associated disease. New Haven, The Connecticut Agricultural Experiment Station.

Stafford, KC, 3rd (1991). "Effectiveness of carbaryl applications for the control of *Ixodes dammini* (Acari: Ixodidae) nymphs in an endemic residential area." J Med Entomol **28**(1): 32-36.

Stafford, KC, 3rd (1993). "Reduced abundance of *Ixodes scapularis* (Acari: Ixodidae) with exclusion of deer by electric fencing." J Med Entomol **30**(6): 986-996.

Stafford, KC, 3rd, Denicola, AJ and Kilpatrick, HJ (2003). "Reduced abundance of *Ixodes scapularis* (Acari: Ixodidae) and the tick parasitoid *Ixodiphagus hookeri* (Hymenoptera: Encyrtidae) with reduction of white-tailed deer." J Med Entomol **40**(5): 642-652.

Stafford, KC, 3rd and Magnarelli, LA (1993). "Spatial and temporal patterns of *Ixodes scapularis* (Acari: Ixodidae) in southeastern Connecticut." J Med Entomol **30**(4): 762-771.

Stanek, G and Reiter, M (2011). "The expanding Lyme Borrelia complex--clinical significance of genomic species?" Clin Microbiol Infect **17**(4): 487-493.

Steere, AC (2001). "Lyme disease." N Engl J Med **345**(2): 115-125.

Steere, AC, Broderick, TF and Malawista, SE (1978). "Erythema chronicum migrans and Lyme arthritis: epidemiologic evidence for a tick vector." Am J Epidemiol **108**(4): 312-321.

Steere, AC, Coburn, J and Glickstein, L (2004). "The emergence of Lyme disease." J Clin Invest **113**(8): 1093-1101.

Steere, AC, Grodzicki, RL, Kornblatt, AN, Craft, JE, et al. (1983). "The spirochetal etiology of Lyme disease." N Engl J Med **308**(13): 733-740.

Steere, AC and Malawista, SE (1979). "Cases of Lyme disease in the United States: locations correlated with distribution of *Ixodes dammini*." Ann Intern Med **91**(5): 730-733.

Steere, AC, Malawista, SE, Bartenhagen, NH, Spieler, PN, et al. (1984). "The clinical spectrum and treatment of Lyme disease." Yale J Biol Med **57**(4): 453-461.

Steere, AC, Malawista, SE, Snydman, DR, Shope, RE, et al. (1977). "Lyme arthritis: an epidemic of oligoarticular arthritis in children and adults in three Connecticut communities." Arthritis Rheum **20**(1): 7-17.

Steere, AC and Sikand, VK (2003). "The presenting manifestations of Lyme disease and the outcomes of treatment." N Engl J Med **348**(24): 2472-2474.

Steere, AC, Sikand, VK, Meurice, F, Parenti, DL, et al. (1998). "Vaccination against Lyme disease with recombinant *Borrelia burgdorferi* outer-surface lipoprotein A with adjuvant. Lyme Disease Vaccine Study Group." N Engl J Med **339**(4): 209-215.

Steere, AC, Sikand, VK, Schoen, RT and Nowakowski, J (2003). "Asymptomatic infection with *Borrelia burgdorferi*." Clin Infect Dis **37**(4): 528-532.

Subramanian, SV and O'Malley, AJ (2010). "Modeling neighborhood effects: the futility of comparing mixed and marginal approaches." Epidemiology **21**(4): 475-478; discussion 479-481.

Talleklint-Eisen, L and Lane, RS (2000). "Spatial and temporal variation in the density of *Ixodes pacificus* (Acari: Ixodidae) nymphs." Environmental Entomology **29**: 272-280.

Telford III, SR (2002). Deer tick-transmitted zoonoses in the Eastern United States. Conservation medicine: ecological health in practice. A Aguirre, RS Ostfeld, GM Tabor, C House and MC Pearl. New York, Oxford University Press**:** 310-324.

Telford III, SR, Mather, TN, Adler, GH and Spielman, A (1990). "Short-tailed shrews as reservoirs of the agents of Lyme disease and human babesiosis." J Parasitol **76**(5): 681-683.

Telford, SR, 3rd, Mather, TN, Moore, SI, Wilson, ML, et al. (1988). "Incompetence of deer as reservoirs of the Lyme disease spirochete." Am J Trop Med Hyg **39**(1): 105-109.

Teutsch, SM and Churchill, RE, Eds. (2000). Principles and Practices of Public Health Surveillance, Oxford.

Tugwell, P, Dennis, DT, Weinstein, A, Wells, G, et al. (1997). "Laboratory evaluation in the diagnosis of Lyme disease." Ann Intern Med **127**(12): 1109-1123.

US Census Bureau. (2012). "State and County QuickFacts." Retrieved 18 September 2012, from http://quickfacts.census.gov/qfd/states/24/24027.html.

Waller, LA, Goodwin, GJ, Wilson, ML, Ostfeld, RS, et al. (2007). "Spatio-temporal patterns in county-level incidence and reporting of Lyme disease in the northeastern United States, 1990-2000." Environ Ecol Stat **14**: 83-100.

Waller, LA and Gotway, CA (2004). Applied Spatial Statistics for Public Health Data. Hoboken NJ, John Wiley and Sons.

Wallis, RC, Brown, SE, Kloter, KO and Main, AJ, Jr. (1978). "Erythema chronicum migrans and Lyme arthritis: field study of ticks." Am J Epidemiol **108**(4): 322-327.

Wey, CL, Griesse, J, Kightlinger, L and Wimberly, MC (2009). "Geographic variability in geocoding success for West Nile virus cases in South Dakota." Health Place **15**(4): 1108-1114.

Wilson, ML (1998). "Distribution and abundance of *Ixodes scapularis* (Acari: Ixodidae) in North America: ecological processes and spatial analysis." J Med Entomol **35**(4): 446-457.

Wilson, ML, Adler, GH and Spielman, A (1985). "Correlation between abundance of deer and that of the deer tick, *Ixodes dammini* (Acari: Ixodidae)." Ann Entomol Soc Am **78**: 172-176.

Wilson, ML, Ducey, AM, Litwin, TS, Gavin, TA, et al. (1990a). "Microgeographic distribution of immature *Ixodes dammini* ticks correlated with that of deer." Med Vet Entomol **4**(2): 151-159.

Wilson, ML, Levine, JF and Spielman, A (1984). "Effect of deer reduction on abundance of the deer tick *(Ixodes dammini)*." Yale J Biol Med **57**(4): 697-705.

Wilson, ML, Litwin, TS, Gavin, TA, Capkanis, MC, et al. (1990b). "Host-dependent differences in feeding and reproduction of *Ixodes dammini* (Acari: Ixodidae)." J Med Entomol **27**(6): 945-954.

Wilson, ML, Telford, SR, 3rd, Piesman, J and Spielman, A (1988). "Reduced abundance of immature *Ixodes dammini* (Acari: Ixodidae) following elimination of deer." J Med Entomol **25**(4): 224-228.

Winters, AM, Eisen, RJ, Delorey, MJ, Fischer, M, et al. (2010). "Spatial risk assessments based on vector-borne disease epidemiologic data: importance of scale for West Nile virus disease in Colorado." Am J Trop Med Hyg **82**(5): 945-953.

Wormser, GP, Dattwyler, RJ, Shapiro, ED, Halperin, JJ, et al. (2006). "The clinical assessment, treatment, and prevention of Lyme disease, human granulocytic anaplasmosis, and babesiosis: clinical practice guidelines by the Infectious Diseases Society of America." Clin Infect Dis **43**(9): 1089-1134.

Yuval, B and Spielman, A (1990). "Duration and regulation of the developmental cycle of *Ixodes dammini* (Acari: Ixodidae)." J Med Entomol **27**(2): 196-201.

Zimmerman, DL, Fang, X and Mazumdar, S (2008). "Spatial clustering of the failure to geocode and its implications for the detection of disease clustering." Stat Med **27**(21): 4254-4266.

Appendix 1.0. Human subjects research approval documentation

Colorado State University
Knowledge to Go Places

Research Integrity & Compliance Review Office
Office of the Vice President for Research
321 General Services Building - Campus Delivery 2011 Fort Collins, CO
TEL: (970) 491-1553
FAX: (970) 491-2293

NOTICE OF APPROVAL FOR HUMAN RESEARCH

DATE:	August 06, 2013
TO:	Peel, Jennifer, 1681 Env & Rad Health Sciences
	Kugeler, Kiersten, 1681 Env & Rad Health Sciences, Nickoloff, Jac, 1681 Env & Rad Health Sciences
FROM:	Barker, Janell, Coordinator, CSU IRB 2
PROTOCOL TITLE:	Assessment of community deer management and Lyme disease incidence
FUNDING SOURCE:	NONE
PROTOCOL NUMBER:	13-4396H
APPROVAL PERIOD:	Approval Date: August 06, 2013 Expiration Date: July 24, 2014

The CSU Institutional Review Board (IRB) for the protection of human subjects has reviewed the protocol entitled: Assessment of community deer management and Lyme disease incidence. The project has been approved for the procedures and subjects described in the protocol. This protocol must be reviewed for renewal on a yearly basis for as long as the research remains active. Should the protocol not be renewed before expiration, all activities must cease until the protocol has been re-reviewed.

If approval did not accompany a proposal when it was submitted to a sponsor, it is the PI's responsibility to provide the sponsor with the approval notice.

This approval is issued under Colorado State University's Federal Wide Assurance 00000647 with the Office for Human Research Protections (OHRP). If you have any questions regarding your obligations under CSU's Assurance, please do not hesitate to contact us.

Please direct any questions about the IRB's actions on this project to:

Janell Barker, Senior IRB Coordinator - (970) 491-1655 Janell.Barker@Colostate.edu
Evelyn Swiss, IRB Coordinator - (970) 491-1381 Evelyn.Swiss@Colostate.edu

Barker, Janell

Janell Barker

Barker, Janell

Approval is to analyze data from all reported cases (~2000 cases) of Lyme disease among residents of Howard County Maryland during 2001-2011.

Approval Period:	August 06, 2013 through July 24, 2014
Review Type:	EXPEDITED
IRB Number:	00000202

STATE OF MARYLAND
DHMH
Maryland Department of Health and Mental Hygiene
INSTITUTIONAL REVIEW BOARD
201 W. Preston Street • Baltimore Maryland 21201
Patricia M. Alt, Ph.D., Chairperson

March 6, 2014

Kiersten Kugeler, MPH
Centers for Disease Control & Prevention
Division of Vector Borne Diseases
3150 Rampart Road
Fort Collins, CO 80521

REF: Protocol #12-10

Dear Ms. Kugeler:

The Maryland Department of Health and Mental Hygiene's Institutional Review Board (IRB) conducted a continuing review of your protocol entitled "Assessment of Community Deer Management and Lyme Disease Incidence." Your protocol was approved through an expedited review process. This approval will expire on **March 16, 2015**. Please refer to the above referenced protocol number in any future modifications or correspondence pertaining to the above named study.

Please be reminded that all of the requirements of the original approval letter remain in effect. Thank you for your continued responsiveness to the IRB requirements and we wish you continued success in your efforts.

If you have any questions, please call the IRB Administrator, Ms. Gay Hutchen. She can be reached at 767-8448.

Sincerely,

Patricia M. Alt

Patricia M. Alt, PhD
Chairperson
Institutional Review Board

cc: IRB Members
Gay Hutchen

410-767-8448 Fax 410-333-7194
Toll Free 1-877-4MD-DHMH TYY for Disabled - Maryland Relay Service 1-800-735-2258
Web Site: www.dhmh.state.md.us/oig/irb

NCEZID Tracking Number: RD-032812KK

Determination of Non-applicability of Human Subjects Regulations

National Center for Emerging and Zoonotic Infectious Diseases (NCEZID)

Project title Assessment of community deer management and Lyme disease incidence

Primary contact Kiersten Kugeler **Division/Branch** DVBD/BDB

The purpose of this form is to document NCEZID's determination that the above-listed protocol does not require submission to CDC's Human Research Protection Office. Under existing institutional policy, authority to determine whether a project is research involving human subjects or whether CDC is engaged in human subjects research is delegated to the National Centers.

Determination

☐ Project does not meet the definition of research under 45 CFR 46.102(d). IRB review is not required.

☐ Project does not involve human subjects under 45 CFR 46.102(f). IRB review is not required.

☒ CDC is not engaged in the conduct of human subjects research per 2008 OHRP engagement guidance. CDC IRB review is not required. *Investigator has provided documentation of appropriate local review.*

Rationale The goal of this study is to provide scientific evidence regarding the relationship of deer reduction in inland settings to human Lyme disease risk. Association between Lyme disease incidence and both hunting status and associated deer density of the census block group will be assessed. Case residence location obtained as part of normal surveillance by the Maryland Department of Health and Mental Hygiene (MDHMH) from Howard County for 2001-2011 will be used to geolocate houses for census block group localization and determine hunting buffer status. Data received at CDC will be broken from any link to the original data at MDHMH that include name, age and sex; but residence location will be known, as it is the critical component of the analysis. The fact that a case occurred at a given residence is not individually identifiable, however; therefore CDC will not be obtaining identifiable information for this study. Since CDC's role in this project will only be the use of unidentifiable data, CDC is not engaged in the work per OHRP guidance and CDC policy applying this guidance to potentially exempt research.

Additional considerations This project was reviewed and approved by the Maryland Department of Health and Mental Hygiene IRB.

Additional requirements: None.

Changes in the nature or scope of this activity may impact the regulatory determination. Please discuss any changes with your NC Human Subjects Advisor before they are implemented.

Reviewed by Wendy Carr, PhD **Title** Human Subjects Advisor, NCEZID

Signature: _____ **Date:** 3-28-2012

PROJECT 1 APPENDICES

Appendix 1.1. Overview of surveillance practices in Howard County

Prior to electronic laboratory reporting, which began in Maryland in 2009, surveillance for human Lyme disease cases in Howard County occurred entirely by paper reports (mostly faxed reports) from clinicians and laboratories. As of 2013, only two laboratories report electronically; others, including some major clinical laboratories, still report by fax (K. Feldman, personal communication). The general dedication of public health employees in Howard County to Lyme disease case follow-up and classification has varied over time. Case counts in Howard County were highest in 2007 and 2008; however, an increase in cases during that time occurred throughout many counties in Maryland. New personnel were hired to work on Lyme disease at the state level in 2006 and 2007; their hiring may have resulted in more robust case follow-up statewide during 2007 and 2008, and resulted in a higher proportion of laboratory reports successfully investigated and classified as cases during those years (K. Feldman, personal communication).

Changes in surveillance practices during 2007-2008 may have led to an increase in proportion of potential cases being classified as confirmed or probable cases. In this scenario, more people with later states of disease (those confirmed by laboratory tests) would be classified as confirmed cases in contrast to earlier years, when those laboratory reports were never successfully classified. Improved case detection and classification during that time would not likely have occurred differentially across space within Howard County. The impact on purely spatial cluster detection was presumably minimal, but impact on temporal analyses could have been more substantial. These changes in surveillance practices underscore the value in

conducting analyses within a single public health jurisdiction to minimize the impact surveillance practices have on spatial variation in disease on a sub-county level.

Appendix 1.2. Assessment of impact of surveillance practices on total case count

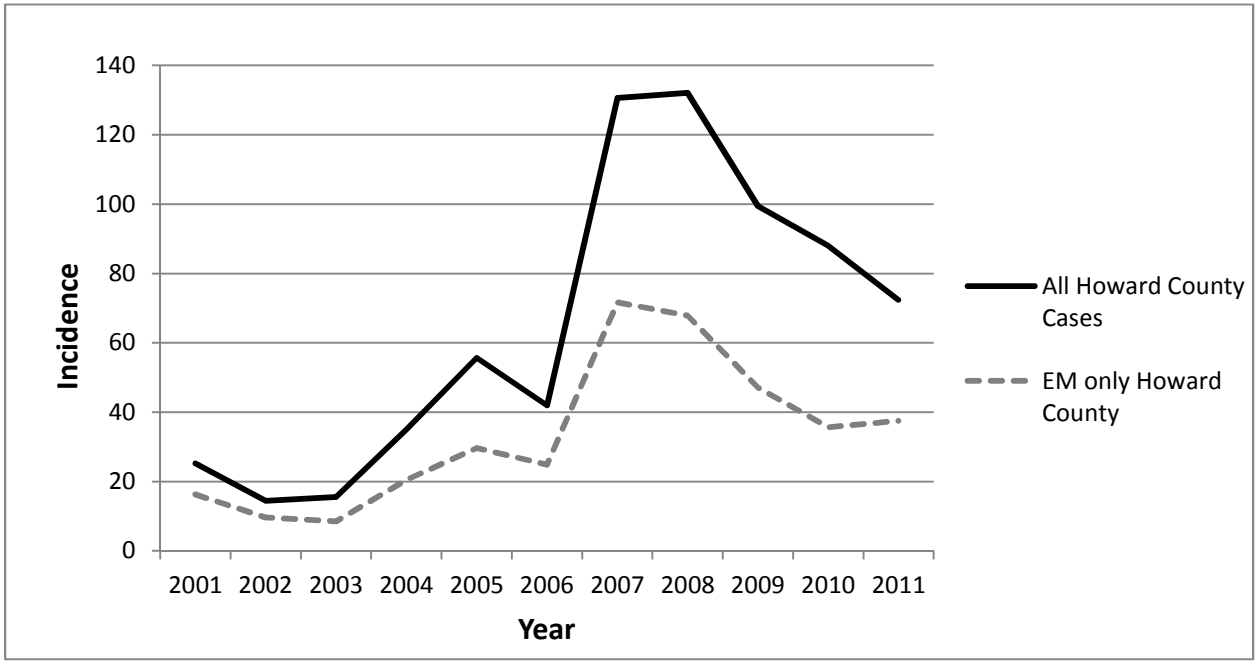


Figure A-1.2. Lyme disease incidence in Howard County 2001-2011: total cases vs. cases with reported erythema migrans (EM)

Erythema migrans (EM) rash is the most common clinical presentation of Lyme disease and the first clinical sign that occurs for most patients, well before an antibody response can be detected. Changes may occur in how laboratory reports are collected (electronic vs. paper) or laboratory criteria required for confirmation. As a result, the proportion of possible cases that can be classified may be modified due to sheer burden on the public health system. In contrast, EM is often reported directly to the public health system by clinicians, a process has not systematically changed over time, and is not subject to changes with modifications to the surveillance case definition that occurred during the study time period. In Figure A-1.2, the same overall trend in incidence during 2001-2011 is detected when examining only cases with

EM as compared to all confirmed and probable cases. The proportion of all cases that have EM rash does, however, decline during the study time period. This trend could be linked to enhanced follow up on laboratory reports, beginning of electronic laboratory reporting as mentioned in Appendix 1.1, and to the beginning of the probable surveillance case definition in 2008.

Appendix 1.3. Spatial autocorrelation of Lyme disease incidence by census block group

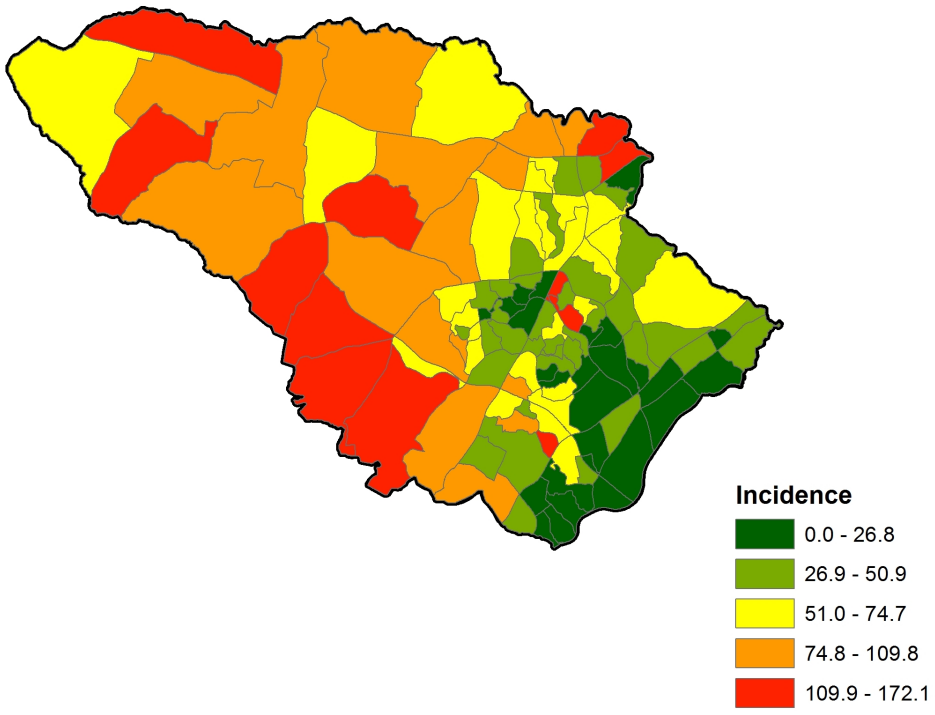


Figure A-1.3a. Cumulative incidence of reported human Lyme disease (confirmed and probable cases per 100,000 residents) by census block group, Howard County, Maryland, 2001-2011

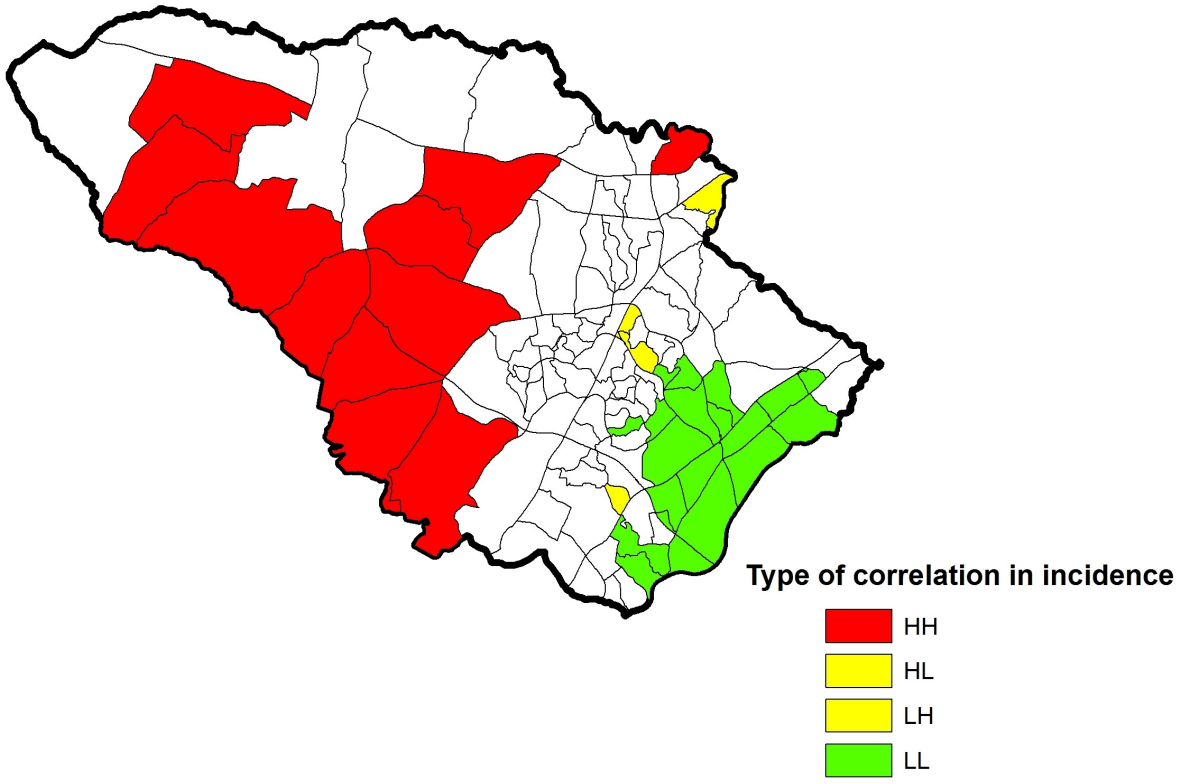


Figure A-1.3b. Census block groups with significant spatial autocorrelation in human Lyme disease incidence. Type of autocorrelation is indicated in the figure legend, with "H" and "L" depicting high and low, respectively.

By examining not only cumulative incidence but also the spatial autocorrelation of cumulative incidence values (by Moran's I test), higher incidence of Lyme disease clearly occurred in the western and central part of Howard County and lower incidence occurred in the eastern part of the County. This appendix presents an alternate method to examine disease clustering than what is presented in the main text of Chapter 2, which was based on binary point-based rather than census block group count data.

Appendix 1.4. Spatial autocorrelation of population growth, 2000-2009

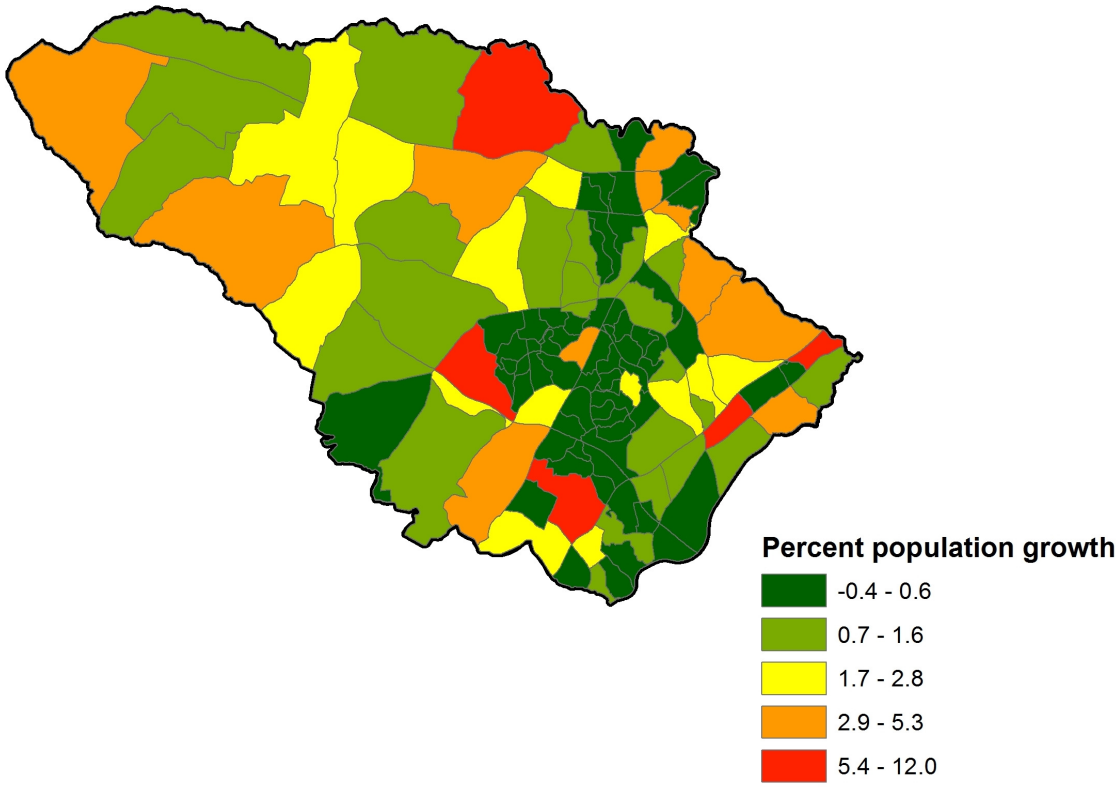


Figure A-1.4a. Percent population growth during 2000-2009 by census block group, Howard County

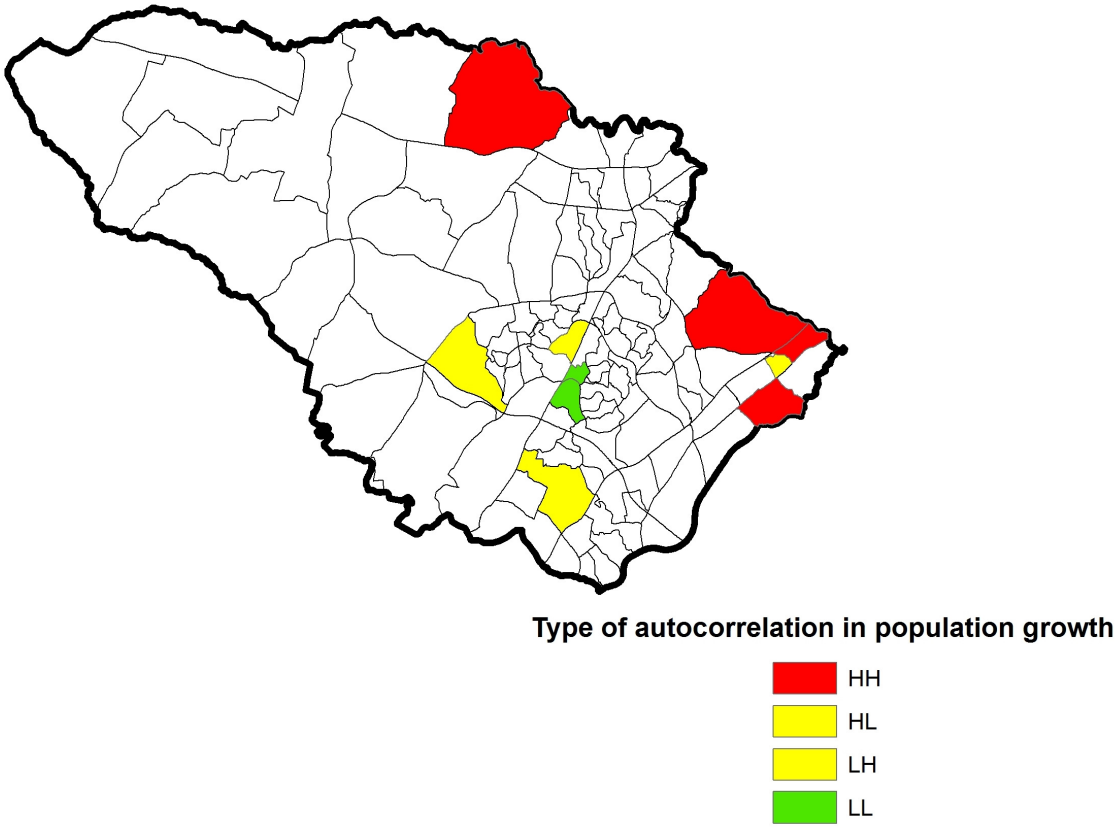


Figure A-1.4b. Census block groups with significant spatial autocorrelation of population growth during 2000-2009. Type of autocorrelation is indicated in the figure legend, with "H" and "L" depicting high and low, respectively.

Differential population growth in the western and central parts of the County during the study period could artificially elevate disease risk in those areas by underestimation of population denominators. The potential of this to affect cluster detection was examined by assessing spatial autocorrelation in estimated population growth from 2000-2009 per census block group. Global Moran's I statistic revealed significant heterogeneity in population growth by census block group ($p<0.001$); local Moran's I statistics revealed clustering of increased population growth in four census block groups (shown in red in Figure A-1.4b.). However, only one of these census block groups was in the elevated risk area of the western and central part

of the County, and was not a block group with very high incidence of Lyme disease. Thus, the detected high-risk clusters do not appear to result from an artificial underestimation of the population.

Appendix 1.5. Spatial cluster detection analysis using case counts in census block groups

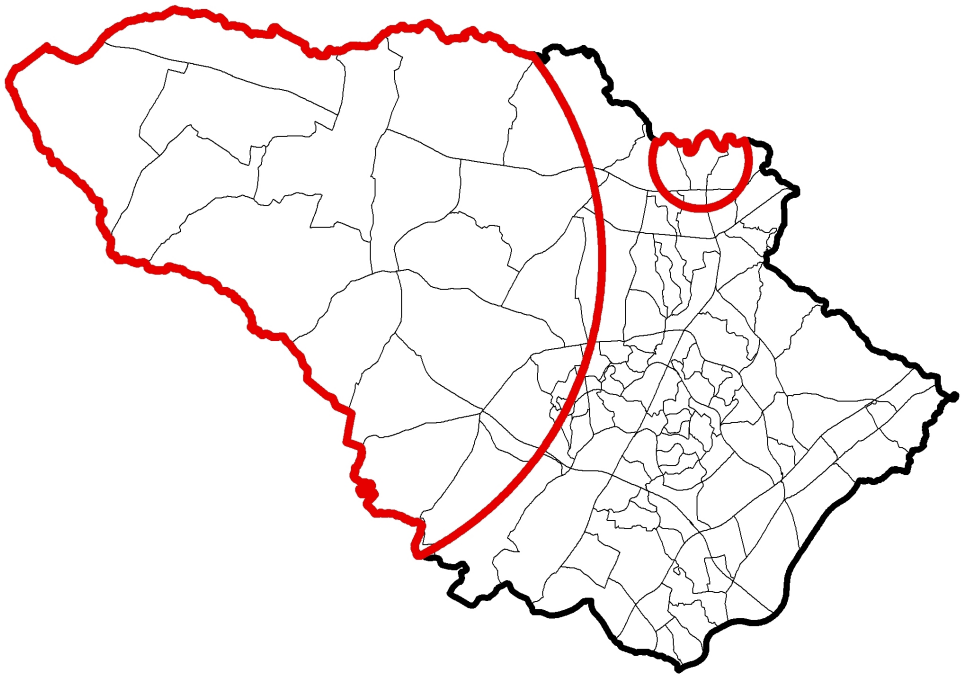


Figure A-1.5. Spatial high-risk clusters based on 50% population maximum cluster size software setting

Cluster findings displayed in A-1.5 were obtained using case counts per census block group rather than binary point-based household data. Although the very small high-risk cluster was not detectable using this method, these results are generally consistent with findings obtained using point-based data.

Appendix 1.6. Spatiotemporal cluster analysis

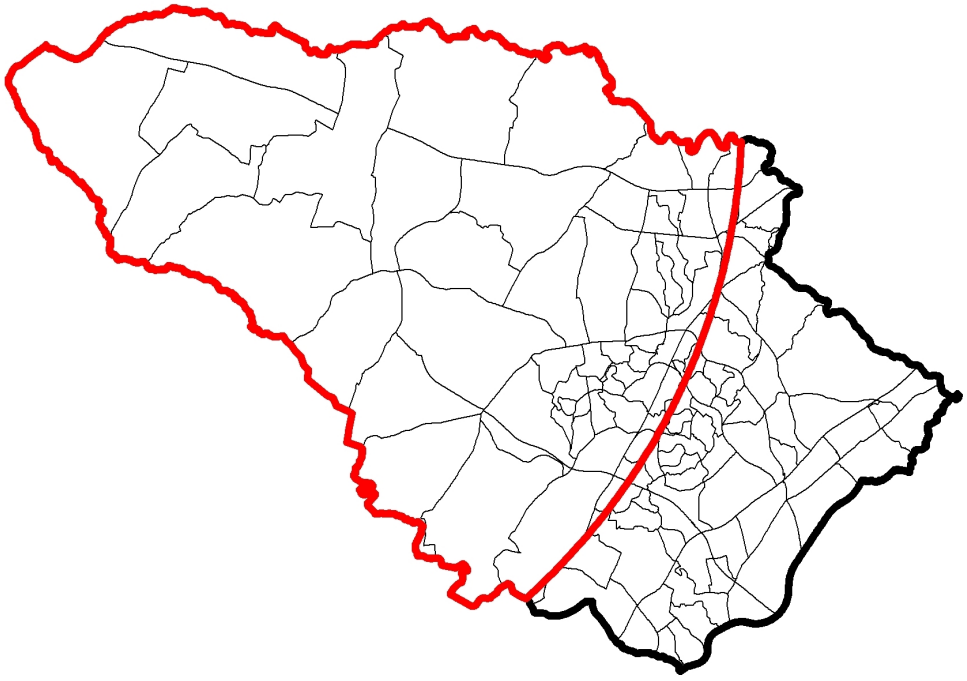


Figure A-1.6. Spatiotemporal cluster analysis with 50% population maximum revealed one high-risk cluster in the western and central part of Howard County beginning June 2005 through end of study period (2011).

Appendix 1.7. Environmental characteristics within clusters assessed at two scales

Table A-1.7. Comparison of land use and land cover characteristics inside and outside high-risk cluster areas assessed at two maximum cluster size settings.

Land use category	**50% population maximum cluster size** (% of total area)		**2 km maximum cluster size** (% of total area)	
	Inside	**Outside**	**Inside**	**Outside**
Low-density development	22.9	10.3	30.3	16.7
Medium-density development	2.4	20.9	13.0	9.8
High-density development	0.7	6.2	1.6	3.1
Urban	2.8	20.1	2.3	10.5
Herbaceous/agricultural	35.0	13.2	13.8	27.0
Forest	28.5	27.7	27.9	23.9
Very low density - forest	5.5	1.7	7.7	3.7
Very low density- agricultural	7.1	1.0	3.1	4.7
Land cover category				
Herbaceous/agricultural	65.4	27.8	35.0	51.1
Urban	3.9	37.7	14.2	18.1
Red-white oak forest	18.2	15.3	38.9	15.3
Chestnut oak forest	5.4	7.1	2.9	6.3
Other deciduous forest	4.7	7.5	4.2	6.0
Mixed forest	1.9	4.0	4.7	2.6
Total forest	30.1	34.0	50.6	30.2

L

and use and land cover characteristics associated with two km maximum high-risk clusters were presented in the main text. Here, we compare those findings to those from high-risk clusters detected using the software default maximum cluster size of 50% of the population. Regardless of maximum cluster size, high-risk clusters had more low-density development, less high-density or urban development, more red and white oak forest but less other deciduous forest. In contrast, those factors that differed depending on the scale of the cluster assessed were: medium-density development, land cover classified as herbaceous or agricultural, and proportion of land classified as forest.

PROJECT 2 APPENDICES

Appendix 2.1. Multilevel Models 2 and 3 with forest fragmentation and socio-demographic variables separately

Table A-2.1a. Multilevel model 2 containing only census block group-level forest indices*

Variable type	Variable	Variable specification** (interquartile range)	Multilevel model		
			aOR†	95% CI‡	p-val
Forest fragmentation indices	Percent classified as forest	20.3%	1.44	1.05-1.97	0.026
	Mean forest patch size	0.15 km^2	1.01	0.88-1.16	0.913
	Forest edge per forest area	1.9 (ratio*1,000)	0.99	0.88-1.16	0.904
	Forest edge per total area	411.5 (ratio*1,000)	0.86	0.72-1.03	0.099

***** Covariance parameter for model: 0.2343, AIC: 16,300

**Associations for continuous variable displayed per increase in interquartile range

†Adjusted odds ratio

‡95% confidence interval

Table A-2.1b. Multilevel model 2 containing only census block group-level socio-demographic indices*

Variable type	Variable	Variable specification** (interquartile range)	Multilevel model		
			aOR†	95% CI‡	p-val
Socio-demographic indices	Population density	2,761 people	0.90	0.77-1.06	0.208
	Percent population growth	2.4%	0.85	0.76-0.95	0.006
	Median home value	$255,600	1.56	1.30-1.88	<0.001
	Median age	8.1 years	0.98	0.82-1.16	0.794
	Percent ≥ bachelor's degree	9.8%	1.14	1.03-1.27	0.014

***** Covariance parameter for model: 0.1328; AIC: 16,481

**Continuous variables displayed per increase in interquartile range

†Adjusted odds ratio

‡95% confideence interval

Table A-2.1c. Multilevel model 3 containing only census block group-level forest indices*

* Covariance parameter for model: 0.1408, AIC: 16,150

Variable type	Variable	Variable specification**	Multilevel model		
			aOR†	95% CI‡	p-val
Land use	Low-density development		2.10	1.73-2.54	<0.001
	Medium-density development		1.95	1.67-2.29	<0.001
	Forest		1.38	1.03-1.84	0.032
	Herbaceous/ agriculture		0.66	0.45-0.95	0.026
	High density development/urban		*ref*		
Land cover	Herbaceous/agriculture		0.96	0.83-1.11	0.567
	Red-white oak forest		1.38	1.17-1.64	<0.001
	Other forest		0.99	0.82-1.19	0.926
	Urban		*ref*		
Forest indices	Percent classified as forest	IQR (16.2%)	1.31	1.00-1.72	0.046
	Mean forest patch size	IQR (0.14 km^2)	1.01	0.90-1.14	0.898
	Forest edge per forest area	IQR (3.74 ratio*1,000)	0.99	0.90-1.14	0.940
	Forest edge per total area	IQR (0.57 ratio*1,000)	0.90	0.77-1.06	0.195

**Referent group for class variables indicated (ref); continuous variables displayed per increase in interquartile range (IQR)
†Adjusted odds ratio
‡95% confidence interval

Table A-2.1d. Multilevel model 3 containing only census block group-level socio-demographic indices*

Variable type	Variable	Variable specification**	Multilevel model		
			aOR†	95% CI‡	p-val
Land use	Low-density development		1.81	1.48-2.20	<0.001
	Medium-density development		1.78	1.52-2.09	<0.001
	Forest		1.28	0.95-1.71	0.010
	Herbaceous/ agriculture		0.61	0.42-0.88	0.009
	High-density development/urban		*ref*		
Land cover	Herbaceous/agriculture		0.91	0.78-1.05	0.194
	Red-white oak forest		1.34	1.13-1.60	0.001
	Other forest		0.98	0.81-1.18	0.834
	Urban		*ref*		
Socio-demographic indices	Population density	IQR (2,761 people)	0.92	0.79-1.08	0.320
	Percent population growth	IQR (2.4%)	0.88	0.79-0.98	0.023
	Median home value	IQR ($255,600)	1.46	1.20-1.76	<0.001
	Median age	IQR (8.1 years)	0.91	0.77-1.07	0.248
	Percent ≥ bachelor's degree	IQR (9.8%)	1.10	1.00-1.22	0.063

*Covariance parameter for model: 0.1098, AIC:16,352
**Referent group for class variables indicated (ref); continuous variables displayed per increase in interquartile range (IQR)
†Adjusted odds ratio
‡95% confidence interval

Appendix 2.2. Red and white oak forest as source for forest fragmentation calculations

Table A-2.2a. Univariate fixed effects red-white oak forest (RWO) and human Lyme disease*

Red and white oak (RWO) forest fragmentation indices	**IQR total**	**Mean (IQR) with Lyme disease**	**Mean (IQR) without Lyme disease**	**Odds ratio** (per IQR increase)	**95% CI**	**p-val**
Percent census block group classified as RWO	0.203	0.174 (0.205)	0.140 (0.195)	1.51	1.40-1.62	<0.001
RWO edge per RWO area (ratio *1000)	1.90	5.325 (1.74)	6.421 (1.90)	0.99	0.99-1.00	0.028
RWO edge per total area (ratio *1000)	411.50	537.670 (368.21)	494.333 (411.50)	1.26	1.18-1.36	<0.001
Mean RWO patch size (km^2)	0.153	0.229406 (0.20249)	0.189749 (0.15256)	1.13	1.10-1.17	<0.001

*Characteristics displayed according to interquartile range (IQR); 95% CI=95% confidence interval

Table A-2.2b. Multilevel Model 2, substituting red and white oak census block group-level forest indices for total forest indices*

Variable type	**Variable**	**Variable specification**** interquartile range	**Multilevel model**		
			aOR†	95% CI‡	p-val
Red-white oak forest (RWO) fragmentation indices	Percent classified as RWO	20.3%	1.39	0.85-2.26	0.189
	Mean RWO patch size	0.15 km^2	1.01	0.87-1.17	0.861
	RWO edge per RWO area	1.9 (ratio*1,000)	1.00	0.99-1.01	0.562
	RWO edge per total area	411.5 (ratio*1,000)	0.90	0.67-1.21	0.477
Socio-demographic indices	Population density	2,761 people	0.95	0.77-1.17	0.636
	Percent population growth	2.4%	0.82	0.73-0.92	0.001
	Median home value	$255,600	1.37	1.11-1.69	0.004
	Median age	8.1 years	1.09	0.90-1.30	0.380
	Percent ≥ bachelor's degree	9.8%	1.14	1.02-1.28	0.027

* Covariance parameter for model: 0.1056, AIC:15,242

**Variables displayed per increase in interquartile range (IQR)

†Adjusted odds ratio

‡95% confidence interval

Table A.2.2c. Multilevel Model 3, substituting red and white oak census block group-level forest indices for total forest indices*

Variable type	Variable	Variable specification**	Multilevel model		
			aOR †	95% CI‡	p-value
Land use	Low-density development		1.90	1.55-2.33	<0.001
	Medium-density development		1.87	1.58-2.21	<0.001
	Forest		1.41	1.04-1.90	0.0264
	Herbaceous/agriculture		0.62	0.42-0.90	0.0132
	High density development/urban		*ref*		
Land cover	Herbaceous/agriculture		0.88	0.75-1.03	0.104
	Red-white oak forest		1.27	1.07-1.52	0.007
	Other forest		0.95	0.78-1.15	0.585
	Urban		*ref*		
Red-white oak forest (RWO) indices	Percent classified as RWO	IQR (20.3%)	1.22	0.80-1.86	0.343
	Mean RWO patch size	IQR (0.15 km^2)	1.05	0.91-1.22	0.509
	RWO edge per RWO area	IQR (1.90 ratio*1,000)	1.00	0.99-1.00	0.471
	RWO edge per total area	IQR (411.5 ratio*1,000)	0.88	0.64-1.20	0.415
Socio-demographic indices	Population density	IQR (2,761 people)	0.98	0.80-1.21	0.863
	Percent population growth	IQR (2.4%)	0.87	0.78-0.98	0.012
	Median home value	IQR ($255,600)	1.30	1.05-1.61	0.016
	Median age	IQR (8.1 years)	1.01	0.85-1.21	0.878
	Percent ≥ bachelor's degree	IQR (9.8%)	1.14	1.02-1.28	0.026

* Covariance parameter for model: 0.0922, AIC: 15117

**Referent group for class variables indicated; continuous variables displayed per increase in interquartile range (IQR)

†Adjusted odds ratio

‡95% confidence interval

Appendix 2.3. Removal of all observations with missing data

By removing all observations with missing land use or land cover, the sample size only decreased by 160 observations. The Model 0 covariance parameter was 0.2944 (as compared to 0.2966). Therefore, these observations had little effect on the initial overall variation across census block groups.

Appendix 2.4. Reclassification of land use categories

Table A-2.4. Results of multilevel Model 3 upon reclassification of very low density development dominated by forest and very low density development dominated by agriculture or open fields from the low density class to the forest/herbaceous classes respectively*

Variable type	Variable	Variable specification**	Multilevel model		
			aOR†	95% CI‡	p-val
Land use	Low-density development		1.71	1.40-2.09	<0.001
	Medium-density development		1.76	1.50-2.06	<0.001
	Forest		1.45	1.12-1.87	0.005
	Herbaceous/agriculture		0.85	0.63-1.15	0.287
	High density development/urban		*ref*		
Land cover	Herbaceous/agriculture		0.91	0.78-1.06	0.223
	Red-white oak forest		1.33	1.12-1.59	0.001
	Other forest		0.97	0.81-1.17	0.762
	Urban		*ref*		
Forest fragmentation indices	Percent classified as forest	IQR (16.2%)	1.32	1.03-1.68	0.959
	Mean forest patch size in CBG	IQR (0.14 km^2)	0.95	0.85-1.07	0.387
	Forest edge per forest area	IQR (3.74 ratio*1,000)	1.03	0.87-1.23	0.725
	Forest edge per total area	IQR (0.57 ratio*1,000)	0.95	0.83-1.10	0.496
Socio-demographic indices	Population density	IQR (2,761 people)	1.01	0.84-1.20	0.959
	Percent population growth	IQR (2.4%)	0.86	0.77-0.95	0.005
	Median home value	IQR ($255,600)	1.51	1.25-1.82	<0.001
	Median age	IQR (8.1 years)	0.98	0.83-1.16	0.817
	Percent ≥ bachelor's degree	IQR (9.8%)	1.10	0.99-1.22	0.086

*Covariance parameter for model: 0.0948, AIC: 16,144
**Referent group for class variables indicated (ref); continuous variables displayed per increase in interquartile range (IQR)
†Adjusted odds ratio
‡95% confidence interval

Here, very low density residential development dominated by agriculture and very low density residential development dominated by forest were reclassified from low-density development into herbaceous/agriculture and forest land use classes, respectively. Both herbaceous/agriculture and forest were land use classes with small sample sizes of households; reclassification yielded larger samples sizes, strengthened the positive association with residence in forest, and attenuated the protective association of residing in land used for herbaceous or agricultural purposes.

Appendix 2.5. Reclassification of land cover categories

Table A-2.5. Results of multilevel Model 3 upon reclassification of land cover categories to more forest categories*

Variable type	Variable	Variable specification**	Multilevel model		
			aOR†	95% CI‡	p-value
Land use	Low-density development		1.82	1.49-2.23	<0.001
	Medium-density development		1.78	1.52-2.09	<0.001
	Forest		1.27	0.94-1.70	0.118
	Herbaceous/agriculture		0.59	0.40-0.85	0.006
	High-density development/urban		*ref*		
Land cover	Herbaceous/agriculture		0.92	0.79-1.07	0.259
	Red-white oak forest		1.33	1.12-1.58	0.001
	Chestnut oak forest		0.95	0.74-1.22	0.681
	Other deciduous forest		0.93	0.69-1.26	0.639
	Mixed deciduous/evergreen forest		1.15	0.82-1.61	0.429
	Urban		*ref*		
Forest fragmentation indices	Percent classified as forest	IQR (16.2%)	1.32	1.04-1.68	0.022
	Mean forest patch size	IQR (0.14 km^2)	0.95	0.85-1.06	0.348
	Forest edge per forest area	IQR (3.74 ratio*1000)	1.03	0.87-1.22	0.739
	Forest edge per total area	IQR (0.57 ratio*1000)	0.96	0.83-1.10	0.507
Socio-demographic indices	Population density	IQR (2,761 people)	1.01	0.85-1.20	0.896
	Percent population growth	IQR (2.4%)	0.87	0.78-0.96	0.008
	Median home value	IQR ($255,600)	1.45	1.21-1.74	<0.001
	Median age	IQR (8.1 years)	0.98	0.83-1.15	0.760
	Percent ≥ bachelor's degree	IQR (9.8%)	1.11	1.00-1.23	0.049

* Covariance parameter for model: 0.08389, AIC: 16,121

**Referent group for class variables indicated (ref); continuous variables displayed per increase in interquartile range (IQR)

†Adjusted odds ratio

‡95% confidence interval

Appendix 2.6. Comparison of multilevel Model 3 with GEE

Table A-2.6. Comparison of multivariable multilevel model 3 of household and census block group risk factors for Lyme disease with model 3 implemented as a generalized estimating equations (GEE) model with an exchangeable correlation structure

Variable type	Variable	Variable specification*	Multilevel model			GEE with exchangeable structure		
			aOR†	95% CI‡	p-val	aOR†	95% CI‡	p-val
Land use	Low-density development		1.85	1.52-2.26	<0.001	1.84	1.51-2.25	<0.001
	Medium-density development		1.80	1.54-2.12	<0.001	1.78	1.47-2.14	<0.001
	Forest		1.29	0.96-1.77	0.089	1.27	0.98-1.65	0.077
	Herbaceous/agriculture		0.62	0.43-0.90	0.012	0.62	0.38-1.01	0.052
	High-density development/ urban		*ref*					
Land cover	Herbaceous/agriculture		0.91	0.78-1.06	0.235	0.92	0.77-1.10	0.343
	Red-white oak forest		1.32	1.11-1.57	0.002	1.33	1.14-1.55	<0.001
	Other forest		0.98	0.81-1.17	0.788	0.98	0.81-1.20	0.857
	Urban		*ref*					
Forest fragmentation indices	Percent classified as forest	IQR (16.2%)	1.36	1.07-1.73	0.012	1.34	1.05-1.71	0.020
	Mean forest patch size	IQR (0.14 km^2)	0.93	0.83-1.04	0.201	0.95	0.87-1.04	0.234
	Forest edge per forest area	IQR (3.74 ratio*1000)	1.03	0.87-1.21	0.772	1.02	0.89-1.17	0.797
	Forest edge per total area	IQR (0.57 ratio*1000)	0.95	0.82-1.09	0.426	0.07	0.83-1.09	0.453
Socio-demographic indices	Population density	IQR (2,761 people)	1.02	0.86-1.22	0.790	1.00	0.84-1.20	0.978
	Percent population growth	IQR (2.4%)	0.87	0.78-0.97	0.010	0.90	0.81-1.00	0.053
	Median home value	IQR ($255,600)	1.46	1.21-1.75	<0.001	1.37	1.11-1.70	0.004
	Median age	IQR (8.1 years)	0.97	0.82-1.15	0.707	0.99	0.82-1.19	0.907
	Percent ≥ bachelor's degree	IQR (9.8%)	1.13	1.01-1.25	0.027	1.11	0.99-1.24	0.084

*Referent group for class variables indicated (ref); continuous variables displayed per increase in interquartile range (IQR)
†Adjusted odds ratio
‡95% confidence interval

PROJECT 3 APPENDICES

Appendix 3.1. Tick counts on deer and deer density using alternate deer density variables

Table A-3.1. Adjusted associations of tick counts on deer during October with mean deer density and variance in deer density across parks, Howard County, Maryland

Exposure variable	Exposure variable specification	Adjusted RR*	95% CI**	p-val
Mean deer density	Continuous (per 10 deer / mi^2)	0.91	0.87-0.96	<0.001
Variance in deer density	Continuous (per 100 units)	0.99	0.99-1.00	<0.001
	Dichotomous low (≤ 810)	*ref*		
	Dichotomous high (> 810)	0.37	0.22-0.60	<0.001

*Adjusted for sex and age of deer and year culled; RR=relative risk; reference group for class variable indicated (ref)

**95% confidence interval

Here, mean and variance in deer density were explored as explanatory factors for association with tick counts on deer. Mean deer density was not divided into levels, as the correlation between mean deer density and the primary explanatory variable, median deer density, was nearly perfect ($r=0.998$, $p<0.001$). The findings were in accordance with those of the primary analysis that used median density.

Variance in deer density in parks had one extreme outlier ($\sigma^2>14{,}000$). When variance was included in adjusted overdispersed Poisson models as a continuous variable, this outlying value was removed. Additionally, variance was dichotomized into low and high values. In this analysis, parks with high variance were inversely associated with tick abundance on deer. Thus, as with the trend in primary analyses for increasing deer density to be associated with fewer ticks on deer, increasing variance in deer density was also associated with fewer ticks on deer. Conversely, those areas with the lowest deer density and least variation in density estimates had the highest tick abundance.

Appendix 3.2. Tick counts on deer and deer density: multinomial and truncated tick counts

Table A-3.2a. Association of tick abundance on deer with deer density in parks, Howard County, Maryland. Tick counts truncated at a high value of 10 ticks

Outcome variable specification	Median deer density level	Adjusted RR*	95 % CI**	p-value
Truncated tick counts (> 10 = 10)	Low (≤ 24 deer per mi^2)	*ref*		
	Medium (25 - 62 deer per mi^2)	1.08	0.80-1.46	0.623
	High (≥ 63 deer per mi^2)	0.50	0.31-0.81	0.004

*Adjusted for sex and age of deer and year culled; reference group for class variable indicated (ref)
**95% confidence interval

The count variable for ticks on deer was truncated at 10; therefore, observations with counts higher than 10 were re-classified to have values of 10. Results in an adjusted overdispersed Poisson model were similar to that of the primary analysis. The highest level of deer density was associated with lowest tick abundance, even when the highest levels of abundance were minimized.

Table A-3.2b. Association of tick abundance on deer with deer density in parks*

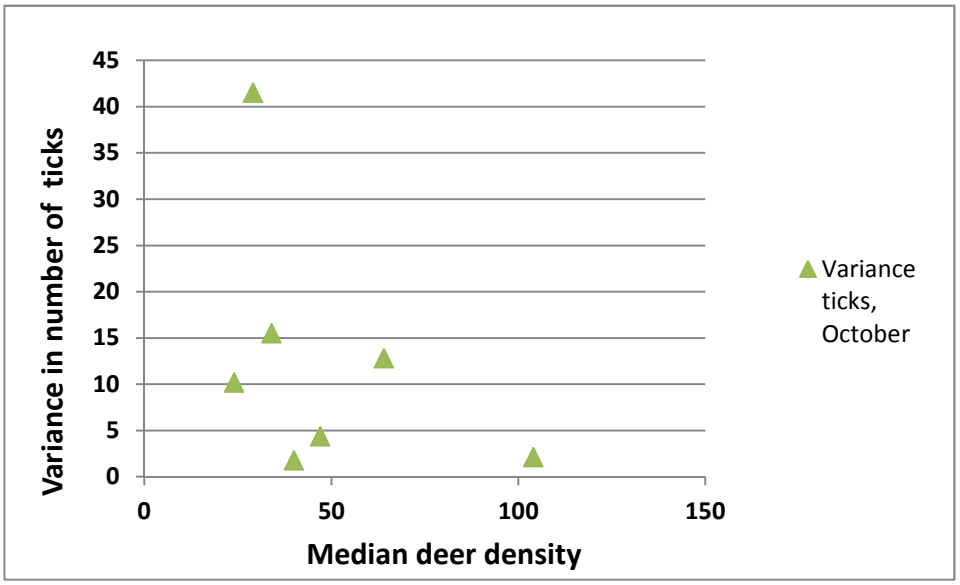
Outcome variable	Median deer density level	aOR†	95% CI‡	p-val
0 ticks	Low (≤ 24 deer per mi^2)	*ref*	*ref*	
1 tick	Medium (25 - 62 deer per mi^2)	1.33	0.78-2.27	0.301
≥ 2 ticks	Medium (25 - 62 deer per mi^2)	1.08	0.69-1.69	0.727
1 tick	High (≥ 63 deer per mi^2)	1.56	0.80-3.02	0.189
≥ 2 ticks	High (≥ 63 deer per mi^2)	0.39	0.20-0.77	0.006

*Tick counts implemented as a three-level multinomial outcome variable
†Adjusted odds ratio; reference group indicated (ref)
‡95% confidence interval

The association between tick counts and deer density was examined with tick counts classified into a three-level nominal variable (no ticks, one tick, and two or more ticks). In this

multinomial adjusted logistic regression model, there were no differences between detection of no ticks vs. one tick in any of the deer density levels. In contrast, the highest deer density level was associated with reduced odds of finding a deer with two or more ticks.

Appendix 3.3. Tick abundance and deer density with alternate variables for tick abundance



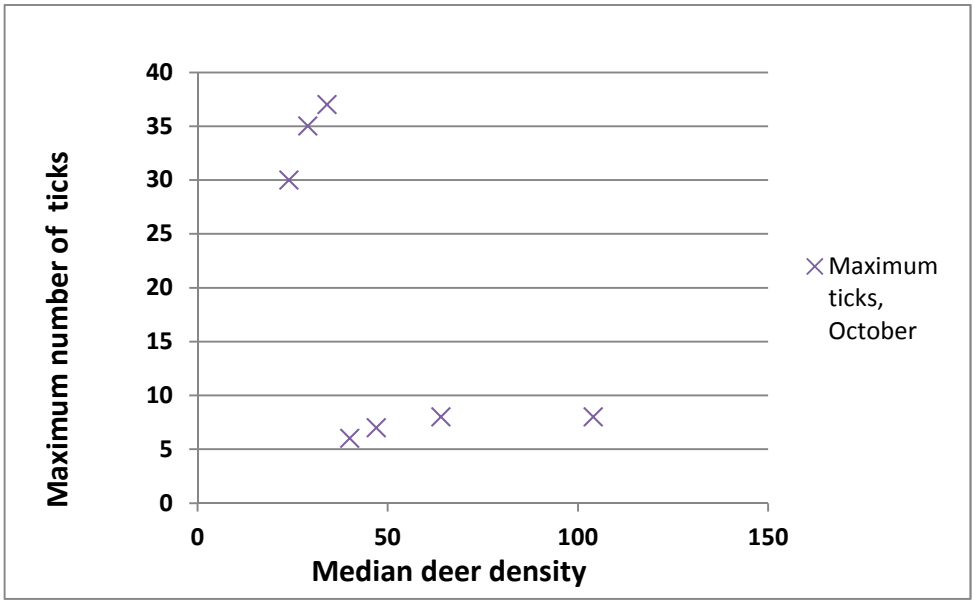


Figure A-3.3. Maximum number of ticks per deer and variance in number of ticks per deer in October according to median deer density during the study period

Adjusted logistic regression models with variance in tick abundance and maximum tick abundance as outcome variables were not created due to quasi-complete or complete separation in data points. Based on visual assessment of this crude association (above), variance in ticks on deer and maximum numbers of ticks on deer were highest at the lowest levels of deer density.

Appendix 3.4. Human Lyme disease outcome models using alternate deer density variables

Table A-3.4. Association of deer density (calculated as mean density and variance in density) and human Lyme disease risk, Howard County, Maryland

	Variable specification	aOR*	95% CI†	p-val	GEE aOR‡	95% CI†	p-val
Mean density	Continuous (10 deer per mi^2)	0.98	0.96-0.99	0.003	0.98	0.96-0.99	<0.001
Quartiles of mean deer density	Q1 (< 41 deer per mi^2)	*ref*			*ref*		
	Q2 (42 - 49 deer per mi^2)	0.93	0.68-1.28	0.667	1.00	0.55-1.82	0.991
	Q3 (50 - 77 deer per mi^2)	0.94	0.66-1.36	0.752	0.96	0.68-1.36	0.824
	Q4 (> 77 deer per mi^2)	0.73	0.56-0.95	0.020	0.70	0.51-0.97	0.030
Variance in density	Low (≤ 810)	*ref*			*ref*		
	High (> 810)	0.63	0.49-0.80	<0.001	0.60	0.44-0.81	0.001

*Adjusted odds ratio; reference group for class variable indicated (ref)
†95% confidence interval
‡Generalized estimating equation (GEE) with exchangeable correlation structure

The primary analysis between deer density and human Lyme disease risk used quartiles of median deer density; there was a trend toward decreased risk of Lyme disease with increased deer density, however the association was not clearly linear. Additionally, mean deer density was examined. The adjusted odds ratio (OR) and adjusted OR in a generalized estimating equations (GEE) model were highly similar to that of median deer density as a continuous variable. Using quartiles of mean deer density, the highest quartile of deer density was associated with reduced risk of Lyme disease, and the dose-response trend according to the fixed effect estimates was clearer than with the quartiles of median deer density.

Appendix 3.5. Association of density of infected adult ticks with human Lyme disease risk

Table A-3.5. Density of infected adult ticks during 2012-2013 and cumulative risk of Lyme disease among households surrounding parks, Howard County, Maryland, 2001-2011*

	OR	95% CI	p-val	aOR	95% CI	p-val
Density of infected adults: 2012-2013	1.01	0.99-1.03	0.351	1.00	0.97-1.02	0.802

*OR=odds ratio; 95% CI= 95% confidence interval; aOR=Adjusted OR

As the density of infected adult ticks increased (in the five parks with corresponding data), no trend toward increase in cumulative human disease risk in those parks was evident. Infection data were only available for the most recent deer hunting season, and did not directly correlate in a biologically meaningful way with human disease data from 2001-2011. Nevertheless, if an association had been detected, it may have suggested that parks with higher burden of infection were constant over time in Howard County.