

# Social vulnerability and Lyme disease incidence: A regional analysis of the United States, 2000-2014

Dhitinut Ratnapradipa <sup>(1)</sup>, Justin Tyler McDaniel <sup>(2)</sup>, Alexandra Barger <sup>(3)</sup>

(1) Department of Population Health, Sam Houston State University, Huntsville, TX

(2) Department of Health Promotion, Charleston Southern University, Charleston, SC

(3) School of Medicine, Southern Illinois University, Carbondale, IL

**CORRESPONDING AUTHOR:** Justin Tyler McDaniel, Department of Health Promotion, Charleston Southern University, [jmcdaniel@csuniv.edu](mailto:jmcdaniel@csuniv.edu)

**DOI:** 10.2427/12158

Accepted on February 13, 2017

## ABSTRACT

**Background:** Lyme disease (LD), which is a highly preventable communicable illness, is the most commonly reported vector borne disease in the USA. The Social Vulnerability Index (SoVI) is a county level measure of SES and vulnerability to environmental hazards or disease outbreaks, but has not yet been used in the study of LD. The purpose of this study was to determine if a relationship existed between the SoVI and LD incidence at the national level and regional division level in the United States between 2000 and 2014.

**Methods:** County level LD data were downloaded from the CDC. County level SoVI were downloaded from the HVRI at the University of South Carolina and the CDC. Data were sorted into regional divisions as per the US Census Bureau and condensed into three time intervals, 2000-2004, 2005-2009, and 2010-2014. QGIS was utilized to visually represent the data. Logarithmic OLS regression models were computed to determine the predictive power of the SoVI in LD incidence rates.

**Results:** LD incidence was greatest in the Northeastern and upper Midwestern regions of the USA. The results of the regression analyses showed that SoVI exhibited a significant quadratic relationship with LD incidence rates at the national level.

**Conclusion:** Our results showed that counties with the highest and lowest social vulnerability were at greatest risk for LD. The SoVI may be a useful risk assessment tool for public health practitioners within the context of LD control.

*Key words:* Lyme disease, social vulnerability, spatial analysis

## INTRODUCTION

Lyme disease (LD) is a tick-borne illness [1-7]. Ticks carrying the LD bacterium are often found on humans in

hard-to-see areas – such as the armpits, groin, or scalp – during the spring and summer months, and require a minimum of 36 to 48 hours to transmit the *B. burgdorferi* bacterium [8]. LD has a low mortality rate, and if it is listed

on a death certificate, is generally only listed as one of several contributing causes [9]; however, it does leave many sufferers with long-lasting multi-organ-system effects, neurological and musculoskeletal symptoms, including difficulty with memory and articulation [10-14], cardiac disease, and Lyme-induced joint inflammation [15, 16]. Zhang et al. [17] estimated that a LD patient, diagnosed with either early- or late-stage disease, accumulates nearly 3,000 USD in direct medical costs. Given that there are approximately 30,000 cases reported to the CDC each year, the yearly economic impact of LD is estimated to be 200 million USD per year [8].

LD is treatable with antibiotics. Treatment, in general, is much more effective if initiated in the early stages of infection [18]. In cases that are not identified, or in which treatment is not initiated until late-stage disease (as is the case in many socially vulnerable/deprived regions in the USA), the course of treatment may be longer, more expensive, and include more intensive therapies [19]. While further research regarding LD diagnostic and treatment protocols is needed, possibly the simplest, least expensive, and most effective way to reduce the burden of LD is to reduce the number of new infections. Personal prevention of LD includes protective behaviors, such as inspecting the body for tick attachment, using insect repellants with DEET or permethrin, and wearing protective clothing, whereas environmental prevention techniques include geographic application of acaricides, control of deer populations, and transformation of tick habitats [20].

Despite its highly preventable nature, LD is the most common vector borne illness in the USA [21] and a significant public health problem [22]. Expansion in the distribution of LD has been observed since the inception of the National Notifiable Diseases Surveillance System (NNDSS) established by the Centers for Disease Control and Prevention (CDC) in 1991 [9]. While the CDC has reported that an average of 30,000 cases of LD occur each year, true incidence – accounting for weaknesses in surveillance and inability of some populations to seek medical treatment – could be “greater than 300,000 cases and as high as one million cases per year in the United States” [23].

According to CDC surveillance data, 95 percent of LD cases in 2013 were reported from only 14 states, each of which is located in one of two regions: the northeastern United States or the midwestern United States [8]. The geographic distribution of LD infection is closely related to the range of the tick vector, and as the tick’s range expands with changing climate, so too does the number of counties reporting a high incidence of LD. The number of high-incidence counties has undergone a remarkable increase in recent decades, expanding outward from the aforementioned geographic centers [9]. The rise and fall of tick populations in a particular region are also associated with fluctuations in the local deer population – which is in turn affected by local food and space availability, predation,

and wildlife management [24]. In addition to understanding these geographic and environmental factors, the study of LD epidemiology might benefit from an understanding of social factors that influence risk of infection. Social vulnerability is one construct that, like LD incidence, has been measured at the county level, and could be used to examine the impact of social factors on LD incidence rates.

Social vulnerability is a measure of a population’s susceptibility to adverse outcomes following a negative event. For example, a socially vulnerable population will suffer more loss and hardship following a natural disaster than a population that is more socially secure. Factors that have been identified to affect social vulnerability include socioeconomic status, gender, age, race and ethnicity, housing status (including urban vs. rural location), occupation, level of education, and access to medical services. More broadly, a higher level of social vulnerability is associated with a lack of political power, low levels of access to resources, limited physical ability, and a lack of social connections. In order to quantify a particular population’s level of vulnerability while accounting for the factors listed above, a composite measure has been developed called the Social Vulnerability Index (SoVI).

Cutter, Boruff, and Shirley [25], authors of the SoVI, examined United States Census data for all 3,141 counties in the United States of America (USA) and, from an initial list of over 250 variables, identified 11 social conditions with principal components factor analysis that explained 76.4 percent of the variance in social vulnerability among counties. Social conditions have been implicated as an upstream contributor to the development of many diseases, both infectious and non-infectious. Studies conducted since the 1960s and 1970s have shown that social indicators are important predictors of longevity and quality of life [25]. The SoVI captures several social and economic variables in a single index, permitting researchers and practitioners the ability to rapidly assess county level socioeconomic status and vulnerability. It is plausible that the SoVI, then, may correlate in some way with LD risk [26], although studies to date have not examined the aforementioned relationship as operationalized in the present study.

## Background

In an effort to control the incidence and distribution of LD in the USA and provide public health professionals with the information needed to prepare their respective communities, several epidemiological studies have been conducted. Numerous risk factors and determinants of LD have been identified in correlational studies. In the present literature review, the following correlates are discussed: climatic factors, regional faunal characteristics, dendrological factors, human knowledge, attitudes, and behavior, encroachment of suburban areas into deciduous forests, and socioeconomic factors.

Early studies showed that ambient temperature patterns, which are characteristic of the Northeastern and Midwestern regions, are hospitable to the survival rates of ticks and thus the incidence of LD [27]. Thus, high temperatures may be a regulating mechanism in tick abundance and the distribution of LD [28]; however, more recent studies have shown that lagged summer temperatures are positively correlated with the reproductive activity of blacklegged ticks and LD incidence, indicating that LD incidence may increase as summer temperatures in the year prior increase [29, 30, 31].

McCabe and Bunnell [32] showed that LD incidence increased linearly with late spring/early summer precipitation, as ticks favor wetter conditions. Other studies have shown that lagged summer precipitation is negatively related with LD incidence, as abundant moisture may result in the proliferation or efficacy of natural tick enemies – proving detrimental to nymphal tick survival [31]. In an effort to resolve these apparent discrepancies, Tran and Waller [33] investigated the divisional affects of precipitation on LD incidence within the Northeastern region of the US and showed that while some regions exhibit a positive correlation (i.e., New England), other regions exhibit a negative correlation (i.e., Mid-Atlantic).

Dendrological variation may influence regional faunal characteristics [34]. This interaction has been hypothesized as a critical determinant of the distribution of LD. The abundance of Oak trees and, thus, the ubiquity of acorns, in the Northeast ensures that populations of vertebrate species on which ticks feed have a life supporting resource [35]. Research has shown that acorn production 1–2 years is strongly and positively correlated with rodent populations [30, 36]. As the community of blacklegged tick hosts increases in density, the likelihood of tick population growth increases [30]. Therefore, acorn crop yields may serve as an important correlate of LD.

The following preventative measures provide protection against LD: avoiding forested areas, wearing protective clothing, performing tick checks, and using insect repellants [20]. Several studies have shown that knowledge of and attitudes towards LD and the aforementioned preventive behaviors are high/appropriate in LD endemic regions [37, 38, 39]. According to the Health Belief Model, knowledge and attitudes are positive determinants of preventive behavior [40]. Thus, residents in endemic areas, theoretically, should possess the motivation to protect themselves against LD; however, adherence to LD preventative behaviors among populations in these areas has not been observed [20].

Low-density suburban sprawl (SS) has been suspected as a risk factor for the development of LD [41]. SS, a phenomenon that has been particularly prevalent in Midwestern and Northeastern divisions, has been described as the migration away from urban areas to and subsequent development of forested areas [42]. Suburban development that fails to preclude significant

forest-herbaceous edges – and only partially fragments woodland areas – permits peridomestic contact with the habitats in which tick populations are characteristically found [43]. Therefore, low-density SS puts residents at risk of being bitten by an infected nymphal tick [44].

In addition to ecological and behavioral factors, socioeconomic status (SES) has been hypothesized as a predictor of LD [45]. Systematic reviews of the literature have shown that SES, often measured as income, plays a significant role in the distribution of LD [46]. Some studies have shown that individuals with higher SES exhibit higher risk for LD than individuals in lower SES classes. For example, Gould et al. [20] revealed that a greater proportion of LD cases were reported in counties with the highest annual median incomes. Cromley and Cromley [47] found similar results.

Other studies, which employed curvilinear modeling techniques, showed that the relationship between SES and LD incidence was quadratic. Specifically, results of two studies have shown that the greatest risk for LD has historically been distributed among those with the least household income and the highest household income [41, 48]. Theoretical reasoning for the aforementioned relationships has been provided in the literature [49, 50].

Social vulnerability has been described as a “lack of access to resources (including information, knowledge, and technology), limited access to political power and representation, social capital ... beliefs and customs ... building stock and age, frail and physically limited individuals, and density of infrastructure” [25, p. 245]. Research has shown that the above vulnerability indicators are often characteristic of particular demographic groups, as defined by gender, race, age, or SES [51, 52, 53]. An index of social vulnerability, called the Social Vulnerability Index (SoVI), was developed by researchers at the Hazards and Vulnerability Research Institute at the University of South Carolina [54] based on the Hazards of Place model [55]. Applications of the SoVI include risk assessment in the incidence of natural or human-caused disasters as well as disease outbreaks [56]; however, hitherto the present paper, the SoVI has only been used in the study of the former [57–64]. Application of the SoVI to LD may afford public health professionals the ability to – with a single planning index – understand their community’s level of risk.

## Purpose

The SoVI provides researchers with a more robust index for understanding SES related risk than has been utilized in previous studies. Because LD symptoms are often hidden – leading to misdiagnosis – the SoVI has the potential to improve health outcomes primary to greater clarity of disease distribution. The purpose of the present study was twofold: (1) to determine if a relationship existed between the SoVI and LD incidence at the national level

and regional division level in the United States between 2000 and 2014, and (2) to determine whether or not the aforementioned relationship, if significant, was consistent across three consecutive five-year time intervals.

## METHODS

Data on annually reported LD cases in the USA between 2000 and 2014 were obtained at the county level from the CDC [21]. Incidence rates per 100,000 were calculated for each county in five-year time intervals, as in Kugeler et al.'s [65] study of Lyme disease, so as to attenuate the effect of county-level transiency and modifications to surveillance techniques: 2000-2004, 2005-2009, and 2010-2014. Incidence rates within each time-interval were standardized based on the standard population of the base year in each interval (i.e., the 2000-2004 time-interval was standardized based on the 2000 standard population, the 2005-2009 time-interval was standardized based on the 2005 standard population, and the 2010-2014 time-interval was standardized based on the 2010 standard population). The aforementioned population data was retrieved from the United States Census Bureau [66]. Secondly, the 2000 and 2005-2009 social vulnerability indices (SoVI) were downloaded for every county in the USA from the HVRI [54], while the 2010 SoVI was downloaded from the CDC [67]. For a more detailed description of the methodology used to produce the SoVI, please refer to Cutter, Boruff, and Shirley [25] and the CDC [67].

Following retrieval of data, we sorted the county level data into nine regional groupings according to the United States Census Bureau's regional division methodology: (1) Northeast – New England; (2) Northeast – Middle Atlantic; (3) Midwest – East North Central; (4) Midwest – West North Central; (5) South – South Atlantic; (6) South – East South Central; (7) South – West South Central; (8) West – Mountain; and (9) West – Pacific [68]. We removed counties from the data set prior to calculation of inferential tests if (a) the incidence rate was zero [69] and (b) the county population was below 100,000 [70].

To the extent that LD incidence rates were skewed in each regional division at each time interval, a Briggsian logarithmic transformation was applied to the LD data [71]. After further study of the logarithmic LD data, we realized – owing to the application of a Kolmogorov-Smirnov test – that the data failed to conform to the Gaussian distribution. To that end – and also because of the common underreporting of LD cases and other weaknesses associated with LD surveillance [21] – all inferential statistical techniques were conducted using bootstrapping with 1,000 resamples [72].

Geometric means of county incidence rates were computed for each USA Census regional division in order to obtain a geographic profile of (a) LD incidence for the

periods 2000-2004, 2005-2009, and 2010-2014 and (b) social vulnerability within each of the same three time periods. QGIS version 2.14 [73] was used to produce visual representations of the means for LD incidence rates and the SoVI at each time interval within each Census division. Pearson correlation coefficients were generated between LD incidence rates and social vulnerability at each time interval so as to manufacture an understanding of the relationship between these two variables at the US Census regional division level. Ordinary least squares (OLS) regression models [74] for each five-year time interval were constructed in an attempt to understand whether or not social vulnerability (SoVI) could predict the incidence of LD. Specifically, the SoVI was the independent variable in each model and LD incidence rates (IR) served as the dependent variable. SoVI and LD datasets were matched spatially and temporally. All models were inspected for a second-degree polynomial given the existence of such relationships in previous studies [41, 48]. Given the Briggsian logarithmic transformation applied to the LD data, the predictive model for each Census division followed the ensuing format:

$$\log_{10}(Y_i) = \alpha + \beta_i X_i + \beta_i X_i^2 + \epsilon_i$$

## RESULTS

Means are shown below for LD incidence rates (Figure 1) and the SoVI (Figure 2) for the periods 2000-2004, 2005-2009, and 2010-2014 for each US Census regional division. The descriptive results showed that LD incidence rates were (a) most prevalent and temporally increasing in the Northeastern and Midwestern regions and (b) lowest in the Western and more centrally located Southern regions across the three time intervals. Expansion of the distribution of LD since 2000 was clearly evident in the Northeastern and Midwestern regions (Figure 1). Social vulnerability decreased in most regions of the USA across the three time intervals, with the exception of the Pacific division in the Western region (Figure 2). The least vulnerable areas in the USA were evident consistently in the northeastern divisions, while the most socially vulnerable geographic location was the West South Central division of the Southern region.

The results of the inferential analysis for 2000-2004 showed that, when aggregated to the national level, social vulnerability had a statistically significant negative relationship with LD incidence,  $F(1, 444) = 25.18$ ,  $p < 0.001$ ,  $R^2 = 0.05$ . At the regional division level, correlation coefficients and beta coefficients for the OLS regression models varied in their representation of the strength of association and prediction of LD incidence rates (Table 1). After Bonferroni adjustment of the per-comparison alpha level, only one model was statistically significant at the

**FIGURE 1.** Lyme disease (LD) incidence rates (IR) per 100,000, aggregated from the county level by US Census regional division for the following 5-year intervals: (A) 2000-2004, (B) 2005-2009, (C) 2010-2014.



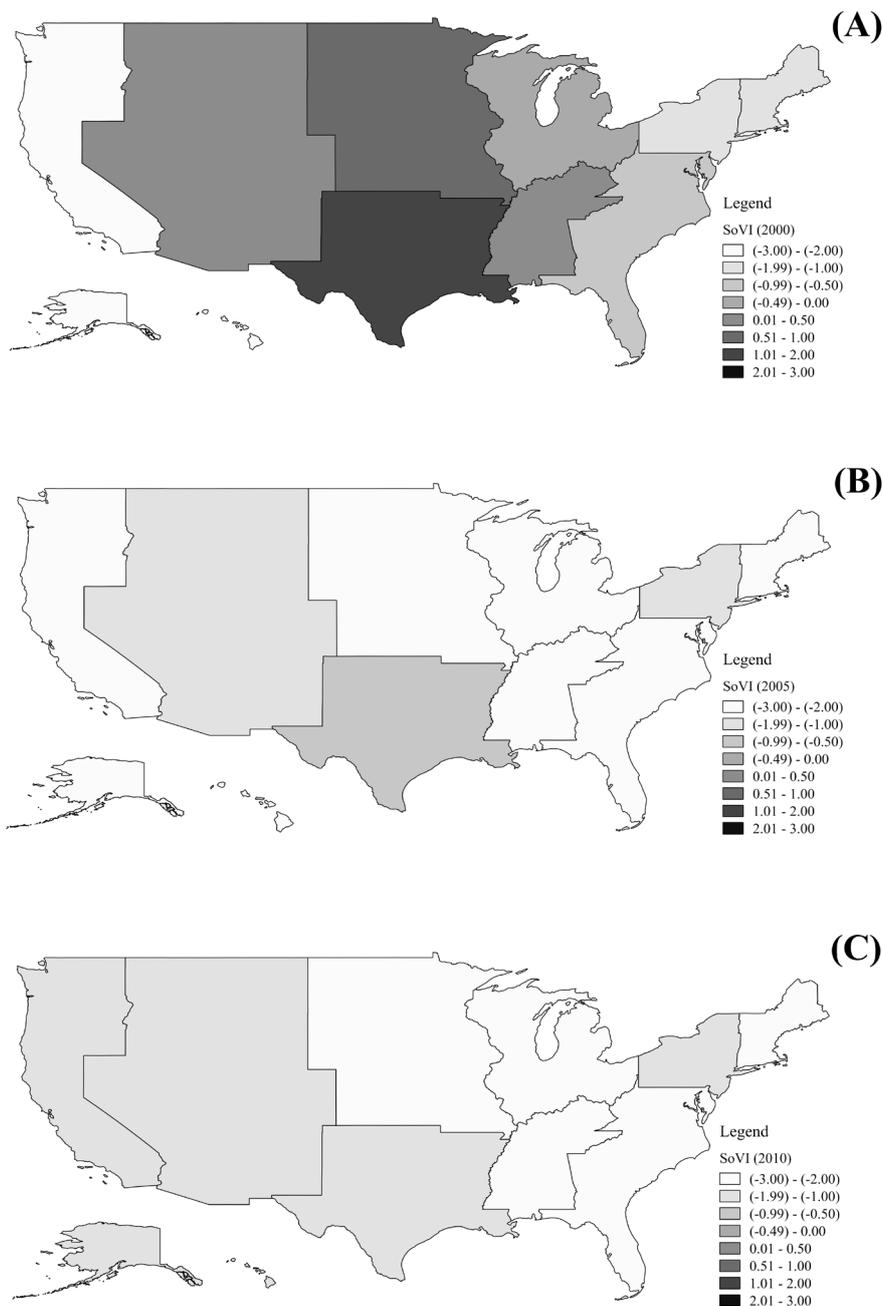
regional division level: the Middle Atlantic (MA) division of the Northeastern region, which exhibited a curvilinear relationship (Table 1). The statistically significant quadratic relationship between the SoVI and LD IR in the MA division indicated that the least and most socially vulnerable counties were at greatest risk for LD.

During the 2005-2009 time interval, social vulnerability – modeled as a quadratic function – explained a small proportion of the variability in LD incidence rates when aggregated to the national level,  $F(2, 477) =$

$27.88, p < 0.001, R^2 = 0.11$ . Table 2 summarizes some of the characteristic divisional differences in the strength of association and prediction of the SoVI for LD incidence. The results obtained from the OLS regression analysis showed that two models were statistically significant after Bonferroni correction to the per-comparison alpha level: the South Atlantic division in the Southern region, which demonstrated a quadratic relationship, and the Middle Atlantic division of the Northeastern region (Table 2).

After analysis of data from the 2010-2014 time

FIGURE 2. Average county SoVI by US Census Divisions for the following time periods; (A) 2000, (B) 2005, and (C) 2010.



interval, a temporal pattern in the association and prediction of LD from/with the SoVI started to emerge. At the national level, the SoVI was a statistically significant quadratic predictor of 2010-2014 county LD incidence,  $F(2, 514) = 23.69, p < 0.001, R^2 = 0.08$  (Table 3). Therefore, data from the most recent time-interval indicated that the relationship between LD and social vulnerability was quadratic – the interpretation of which follows: the greatest and least socially vulnerability counties were at greater risk for LD than counties with less polarized SoVI scores. Similarly to the 2005-2009 time interval, only

two Census divisions emerged with statistically significant regression models in the 2010-2014 time interval: the South Atlantic division in the Southern region, which exhibited a quadratic function, and the Middle Atlantic division of the Northeastern region.

## DISCUSSION

The purpose of the present study was twofold: (1) to determine if a relationship existed between the SoVI and LD

**TABLE 1. OLS Regression Models for Lyme disease by US Census Region Divisions (2000-2004).**

		B	SE B	BCa 95% CI		F	df	p		r	R <sup>2</sup> <sup>a</sup>
				Lower	Upper						
<b>All USA Counties</b>											
	Intercept	0.69	0.04	0.61	0.77	25.18	(1, 444)	0.001	*		0.05
	SoVI	-0.06	0.01	-0.89	-0.04					-0.23	
<b>Western Region</b>											
Pacific	Intercept	0.08	0.09	-0.08	0.23	0.81	(1, 47)	0.37			0.02
	SoVI	-0.01	0.01	-0.04	0.01					-0.13	
Mountain	Intercept	-0.14	0.07	-0.28	0.04	0.94	(1, 11)	0.35			0.08
	SoVI	0.05	0.07	-0.03	0.20					0.28	
<b>Midwestern Region</b>											
West North Central	Intercept	0.73	0.15	0.43	1.01	2.34	(1, 29)	0.14			0.08
	SoVI	-0.08	0.06	-0.17	0.04					-0.27	
East North Central	Intercept	0.40	0.06	0.30	0.52	0.27	(1, 83)	0.60			
	SoVI	-0.01	0.02	-0.06	0.03					-0.06	0.01
<b>Southern Region</b>											
West South Central	Intercept	0.14	0.07	-0.01	0.28	3.24	(1, 37)	0.08			0.08
	SoVI	-0.04	0.02	-0.08	-0.01					-0.28	
East South Central	Intercept	0.31	0.05	-0.21	0.42	0.13	(1, 19)	0.73			0.01
	SoVI	-0.01	0.02	-0.04	0.02					-0.08	
South Atlantic	Intercept	0.54	0.06	0.41	0.68	6.47	(1, 95)	0.01			0.06
	SoVI	-0.06	0.02	-0.10	-0.01					-0.25	
<b>Northeastern Region</b>											
Middle Atlantic	Intercept	1.69	0.08	1.52	1.87	11.55	(2, 74)	0.001	*		0.24
	SoVI	-0.15	0.03	-0.20	-0.09					-0.48	
	SoVI 2	-0.01	0.01	-0.02	-0.01					-0.33	
New England	Intercept	2.08	0.21	1.76	2.55	5.22	(1, 11)	0.04			0.32
	SoVI	-0.16	0.12	-0.39	-0.07					-0.57	

Note. SoVI was used as the predictor variable for each model. Bootstrap results are based on 1,000 bootstrap samples. All regression models are based on a natural logarithmic transformation of the Lyme disease incidence rates from 2000-2004. Bonferroni adjusted alpha value = 0.005. <sup>a</sup> Quadratic models shown above exhibited statistically significant (< 0.05) changes in R<sup>2</sup> values when the squared term was entered into the model.

incidence at the national level and regional division level in the United States between 2000 and 2014, and (2) to determine whether or not the aforementioned relationship, if significant, was consistent across three consecutive five-year time intervals. Although the total explained variance in our models was relatively low, our results showed that the SoVI could be used in the prediction of LD incidence. At the national level, the SoVI predicted a maximum of 11 percent of the variance in LD. The national models for each time interval, with the exception of the 2000-2004 time-interval, were best fitted with a second-degree polynomial, indicating that counties with the lowest and the highest social vulnerability exhibited the greatest risk for LD.

One regional division exhibited statistically

significant regression models at each of the three time intervals: the Middle Atlantic division. During the 2000-2004 time-interval, the aforementioned relationship was quadratic, while in the two more proximal time-intervals the aforementioned relationship failed to reach statistically significant changes in R<sup>2</sup> with the inclusion of a second-degree polynomial. In the 2005-2009 and the 2010-2014 time-intervals, the South Atlantic division exhibited significant quadratic relationships – indicating that the greatest and least socially vulnerable counties in this division were at greatest risk for LD. A greater proportion of the variance in LD was accounted for in the regional division models than the national models. Specifically, while the greatest proportion of

**TABLE 2. OLS Regression Models for Lyme disease by US Census Region Divisions (2005-2009).**

		B	SE B	BCa 95% CI		F	df	p		r	R <sup>2</sup> <sup>a</sup>
				Lower	Upper						
<b>All USA Counties</b>											
	Intercept	0.59	0.05	0.49	0.68	27.88	(2, 477)	< 0.001	*		0.11
	SoVI	-0.07	0.02	-0.12	-0.04					-0.14	
	SoVI 2	0.01	0.02	0.01	0.02					0.11	
<b>Western Region</b>											
Pacific	Intercept	0.16	0.11	-0.05	0.37	0.025	(1, 47)	0.88			0.01
	SoVI	0.01	0.03	-0.06	0.07					0.02	
Mountain	Intercept	0.17	0.08	0.02	0.33	0.29	(1, 21)	0.60			0.01
	SoVI	0.02	0.03	-0.51	0.07					0.12	
<b>Midwestern Region</b>											
West North Central	Intercept	0.36	0.22	-0.03	0.82	9.40	(1, 28)	0.01			0.25
	SoVI	-0.19	0.05	-0.29	-0.07					-0.50	
East North Central	Intercept	0.41	0.07	0.29	0.54	5.60	(1, 93)	0.02			0.06
	SoVI	-0.07	0.03	-0.13	-0.02					-0.24	
<b>Southern Region</b>											
West South Central	Intercept	0.14	0.07	0.12	0.24	8.06	(1, 40)	0.01			0.17
	SoVI	-0.06	0.02	-0.12	-0.03					-0.41	
East South Central	Intercept	-0.01	0.13	-0.26	0.21	2.13	(1, 19)	0.16			0.10
	SoVI	-0.06	0.06	-0.16	0.02					-0.32	
South Atlantic	Intercept	0.37	0.06	0.25	0.49	26.94	(2, 104)	< 0.001	*		0.34
	SoVI	0.02	0.04	-0.07	0.11					-0.50	
	SoVI 2	0.02	0.01	0.01	0.04					0.58	
<b>Northeastern Region</b>											
Middle Atlantic	Intercept	1.70	0.08	1.52	1.85	31.51	(1, 74)	< 0.001	*		0.30
	SoVI	-0.13	0.02	-0.17	-0.10					-0.55	
New England	Intercept	2.11	0.15	1.82	2.44	7.68	(1, 30)	0.01			0.20
	SoVI	-0.09	0.04	-0.18	-0.01					-0.45	

Note. SoVI was used as the predictor variable for each model. Bootstrap results are based on 1,000 bootstrap samples. All regression models are based on a natural logarithmic transformation of the Lyme disease incidence rates from 2005-2009. Bonferroni adjusted alpha value = 0.005.

<sup>a</sup> Quadratic models shown above exhibited statistically significant (< 0.05) changes in R<sup>2</sup> values when the squared term was entered into the model.

explained variance in LD at the national level was 11 percent, the regional division models demonstrated far greater explanatory power. Specifically, the statistically significant regional division models explained at least 24 percent of the variance in LD.

While the explained variance in each of the aforementioned models was low, the usefulness of the SoVI as a predictor of Lyme disease should not be dismissed for two primary reasons. First, as discussed in Abelson [75], models with low explained variance do not excuse their respective independent variables from significant influence on the dependent variable

– especially in cases where the index grows over time, as was demonstrated between 2000-2004 and 2010-2014 in the present study. Specifically, the R<sup>2</sup> value for the national models increased by 0.03 from 2000-2004 to 2010-2014. Second, the present study is exploratory. The importance of socioeconomic status and vulnerability has been demonstrated; however, the inclusion of additional variables will be needed in future studies in order to comprehensively and more robustly estimate LD incidence at the county level. Candidate variables for future studies are discussed momentarily.

**TABLE 3. OLS Regression Models for Lyme disease by US Census Region Divisions (2010-2014).**

		B	SE B	BCa 95% CI		F	df	p	r	R <sup>2</sup> <sup>a</sup>
				Lower	Upper					
<b>All USA Counties</b>										
	Intercept	0.67	0.05	0.57	0.78	23.69	(2, 514)	0.001	*	0.08
	SoVI	-0.06	0.02	-0.12	-0.01					-0.16
	SoVI 2	0.01	0.01	0.01	0.02					0.27
<b>Western Region</b>										
Pacific	Intercept	0.14	0.11	-0.06	0.04	0.50	(1, 53)	0.50		0.01
	SoVI	-0.03	0.04	-0.10	0.06					-0.10
Mountain	Intercept	0.17	0.07	0.04	0.30	0.39	(1, 25)	0.54		0.02
	SoVI	0.02	0.03	-0.04	0.08					0.12
<b>Midwestern Region</b>										
West North Central	Intercept	0.67	0.05	0.10	1.29	3.43	(1, 30)	0.07		0.10
	SoVI	-0.15	0.10	-0.31	0.06					-0.32
East North Central	Intercept	0.52	0.08	0.38	0.67	7.99	(1, 97)	0.01		0.08
	SoVI	-0.09	0.03	-0.16	-0.04					-0.28
<b>Southern Region</b>										
West South Central	Intercept	-0.03	0.06	-0.16	0.11	7.34	(1, 36)	0.01		0.17
	SoVI	-0.06	0.02	-0.10	-0.02					-0.41
East South Central	Intercept	0.15	0.14	-0.12	0.57	1.34	(1, 26)	0.26		0.05
	SoVI	-0.05	0.05	-0.12	0.11					-0.22
South Atlantic	Intercept	0.56	0.06	0.45	0.69	19.87	(2, 121)	0.001	*	0.25
	SoVI	0.02	0.04	-0.06	0.11					-0.41
	SoVI 2	0.03	0.01	0.01	0.04					0.50
<b>Northeastern Region</b>										
Middle Atlantic	Intercept	1.96	0.05	1.86	2.06	24.12	(1, 76)	0.001	*	0.24
	SoVI	-0.09	0.02	-0.12	-0.05					-0.49
New England	Intercept	2.31	0.10	2.13	2.58	7.48	(1, 30)	0.01		0.20
	SoVI	-0.07	0.03	-0.11	0.01					-0.45

Note. SoVI was used as the predictor variable for each model. Bootstrap results are based on 1,000 bootstrap samples. All regression models are based on a natural logarithmic transformation of the Lyme disease incidence rates from 2010-2014. Bonferroni adjusted alpha value = 0.005.

<sup>a</sup> Quadratic models shown above exhibited statistically significant (< 0.05) changes in R<sup>2</sup> values when the squared term was entered into the model.

### Agreement with Previous Research

Our results were consistent with previous studies on the relationship between social indicators and LD incidence [41, 48]. To the extent that our analysis indicated that the relationship between SoVI and LD incidence was – primarily – quadratic, especially with regard to the most recently available data from the CDC, public health efforts should be directed accordingly. In particular, given the expanding distribution of LD in the Northeastern and Midwestern regions [9], focus should be directed towards highly and minimally vulnerable counties in these areas.

The SoVI in this case highlights possible environmental and social differences that influence variability in disease risk between geographically separated populations. The SoVI may be useful as a tool for identifying regions and target populations within which public health education will be most effective at reducing the incidence of disease. In areas where LD is prevalent, state and local health departments could host awareness campaigns to disseminate LD prevention information. Healthcare providers in these areas should be more vigilant about testing for LD, making early detection and treatment more likely. The combination of prevention, early detection, and

treatment – afforded by awareness of a heightened risk of LD in a particular region or population – should reduce LD associated morbidity and economic loss.

### Limitations and Future Directions

This study is the first to investigate the relationship between social vulnerability, as measured by the SoVI [54], and LD incidence in the USA. While our analysis yielded consistently significant results at three time intervals, the regression models presented herein should be interpreted with caution owing to the presence of several limitations. First, our analysis was limited in that it was conducted with county-level, population data. Second, to the extent that LD cases may be underreported [24], our geographic depiction of the distribution of LD may have been conservative.

Third, we did not account for population transience. Fourth, the scope of the present study was limited to the relationship between social vulnerability and LD incidence. Because we only included one independent variable in our analysis, the total explained variance in our regression models was low. Several other variables have been shown to exert influence on the development of LD. Future studies should be conducted with a more comprehensive focus at finer levels of resolution – that is, future studies should examine the influence of weather, forestation, and, perhaps, social media influence, in addition to social vulnerability. The World Health Organization has described several areas internationally, besides the USA, that provide a hospitable environment for the development of Lyme disease [76]: Asia, central, north-western, and eastern Europe. The present study provided an investigation of the USA only; therefore, future studies should consider the relationship between social vulnerability and LD in Asia and Europe.

Given the increasing distribution of Lyme disease, coupled with the fact that LD is the most common vector-borne disease in the USA, policy-level efforts directed toward the prevention of LD are needed. First, research is needed for the development of an effective LD vaccine, as such an intervention could provide the greatest population health benefits, especially for residents in the greatest and least vulnerable counties in the USA [77]. Second, the promotion of entomological based approaches to prevention, including the reduction of the LD vector, could minimize the national LD burden. Third, more directed efforts are needed toward ensuring the equal dissemination of benefits from legislation developed to create a tick-borne advisory committee in the USA [78]. Fourth, because policy and planning efforts among government officials often occur at the county and census tract level, the SoVI provides a useful tool for Lyme disease risk assessment [59]. The SoVI illustrates that vulnerability and, in particular, capacity for responsiveness to disease

outbreaks (Lyme disease specifically) is not uniform across counties in the United States. Resources – such as those that might advance the second policy implication referenced above – should be distributed and prioritized based on an understanding of the SoVI, as such efforts will minimize the risk for the expansion of human Lyme disease.

### Acknowledgments

This paper was presented in Washington, DC, at the 37<sup>th</sup> annual meeting of the Society of Behavioral Medicine on April 1, 2016.

### Competing interests statement

The authors have no conflicts of interest to report.

### REFERENCES

1. Burgdorfer W, Barbour AG, Hayes SF, Benach JL, Grunwaldt E, Davis JP. Lyme disease: A tick-borne spirochetosis? *Science* 1982; 216: 1317-9.
2. Steere AC, Grodzicki RL, Kornblatt AN, et al. The spirochetal etiology of Lyme disease. *New Engl J Med* 1983; 308(13): 733-40.
3. Stanek G, Wormser GP, Gray J, Strle F. Lyme borreliosis. *Lancet* 2012; 379: 461-473. doi: 10.1016/S0140-6736(11)60103-7
4. Anderson JF. Mammalian and avian reservoirs for *Borrelia burgdorferi*. *Ann NY Acad Sci* 1988; 539: 180-91.
5. Brown RN, Lane RS. Lyme disease in California: A novel enzootic transmission cycle of *Borrelia burgdorferi*. *Science* 1992; 256: 1439-42.
6. Piesman J. Transmission of Lyme disease spirochetes (*Borrelia burgdorferi*). *Exp Appl Acarol* 1989; 7: 71-80.
7. Wood CL, Lafferty KD. Biodiversity and disease: A synthesis of ecological perspectives on Lyme disease transmission. *Trends Ecol Evol* 2012; 28(4): 239-47.
8. Centers for Disease Control and Prevention. Lyme disease. Retrieved from: <http://www.cdc.gov/lyme/transmission/index.html> [Accessed January 12, 2016].
9. Kugeler KJ, Farley GM, Forrester JD, Mead PS. Geographic distribution and expansion of human Lyme disease, United States. *Emerg Infect Dis* 2015; 21(8): 1455-7. doi: 10.3201/eid2108.141878
10. Benke T, Gasse T, Hittmair-Delazer M, Schmutzhard E. Lyme encephalopathy: long-term neuropsychological deficits years after acute neuroborreliosis. *Acta Neurol Scand* 2009; 91(5): 353-7.
11. Duray PH, Steere AC. Clinical pathologic correlations of Lyme disease by stage. *Ann NY Acad Sci* 2006; 539: 65-79.
12. Feder HM, Johnson BJB, O'Connell S, Shapiro ED, Steere AC, Wormser GP. A critical appraisal of chronic Lyme disease. *New Eng J Med* 2007; 357(14): 1422-30.
13. Logigian EL, Kaplan RF, Steere AC. Chronic neurologic manifestations of Lyme disease. *New Eng J Med* 2009; 323: 1438-44.
14. Steere AC. Lyme disease. *New Eng J Med* 1989; 321: 586-96.
15. Steere AC, Batsford WP, Weinberg M, et al. Lyme Carditis:

- Cardiac abnormalities of Lyme disease. *Ann Intern Med* 1980; 93(1): 8-16.
16. Steere AC, Schoen RT, Taylor E. The clinical evolution of Lyme arthritis. *Ann Intern Med* 1987; 107(5): 725-31.
  17. Zhang X, Meltzer MI, Pena CA, Hopkins AB, Wroth L, Fix AD. Economic impact of Lyme disease. *Emerg Inf Dis* 2006; 12(4): 653-60.
  18. Wormser GP, Dattwyler RJ, Shapiro ED, et al. 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 Inf Dis* 2006; 43(9): 1089-134.
  19. Steere AC, Sikand VK. The presenting manifestations of Lyme disease and the outcome of treatment (letter to the editor). *New Eng J Med* 2003; 348: 2472-4.
  20. Gould LH, Nelson RS, Griffith KS, et al. Knowledge, attitudes, and behaviors regarding Lyme disease prevention among Connecticut residents, 1999-2004. *Vector-Borne Zoonot* 2008; 8(6): 769-775. doi: 10.1089/vbz.2007.0221
  21. Centers for Disease Control and Prevention. Lyme disease: Data and statistics. Retrieved from: <http://www.cdc.gov/lyme/stats/> [Accessed December 14, 2016].
  22. Perronne C. Lyme and associated tick-borne diseases: Global challenges in the context of a public health threat. *Front Cell Inf Microb* 2014; 4(74): 1-6. doi: 10.3389/fcimb.2014.00074
  23. Stricker RB, Johnson L. Lyme disease: Call for a Manhattan Project to combat the epidemic. *PLOS Path* 2014; 10(1): 1-3. doi: 10.1371/journal.ppat.1003796
  24. Duffy DC, Campbell SR, Clark D, DiMotta C, Gurney S. Ixodes scapularis (Acari: Ixodidae) deer tick Mesoscale populations in natural areas: Effects of deer, area, and location. *J Med Entomol* 1994; 31(1): 152-8.
  25. Cutter SL, Boruff BJ, Shirley WL. Social vulnerability to environmental hazards. *Soc Sci Quart* 2003; 84(2): 242-61.
  26. Link BG, Phelan J. Social conditions as fundamental causes of disease. *J Health Soc Behav* 1995; 35: 80-94.
  27. Ashley ST, Meentemeyer V. Climatic analysis of Lyme disease in the United States. *Clim Res* 2004; 27: 177-87.
  28. Mount GA, Haile DG, Daniels E. Simulation of blacklegged tick (Acari: Ixodidae) population dynamics and transmission of *Borrelia burgdorferi*. *J Med Entomol* 1997; 34: 461-84.
  29. Ogden NH, Radojevic M, Wu X, Duwuri VR, Leighton PA, Wu J. Estimated effects of projected climate change on the basic reproductive number of the Lyme disease vector *Ixodes scapularis*. *Environ Health Persp* 2014; 122(6): 631-8.
  30. Ostfeld RS, Canham CD, Oggenfuss K, Winchcombe RJ, Keesing F. Climate, deer, rodents, and acorns as determinants of variation in Lyme disease risk. *PLoS Bio* 2006; 4(6): e145.
  31. Schaub EM, Ostfeld RS, Evans A. What is the best predictor of annual Lyme disease incidence: Weather, mice, or acorns? *Ecol App* 2005; 15: 575-86.
  32. McCabe GJ, Bunnell JE. Precipitation and the occurrence of Lyme disease in the Northeastern United States. *Vector-Borne Zoonot* 2004; 4(2): 143-8.
  33. Tran P, Waller L. Variability in results from negative binomial models for Lyme disease measured at different spatial scales. *Environ Res* 2015; 136: 373-80. doi: 10.1016/j.envres.2014.08.041
  34. Goodrum PD, Reid VH, Boyd CE. Acorn yields, characteristics, and management criteria of oaks for wildlife. *J Wildlife Manage* 1971; 35(3): 520-532. doi: 10.2307/3799707
  35. Ostfeld RS, Jones CG, Wolff JO. Of mice and mast: Ecological connections in eastern deciduous forests. *Biosci* 1996; 46: 323-30.
  36. Ostfeld RS, Schaub EM, Canham CD, Keesing F, Jones CG, Wolff JO. Effects of acorn production and mouse abundance on abundance and *Borrelia burgdorferi* infection prevalence of nymphal Ixodes scapularis ticks. *Vector-Borne Zoonot* 2001; 1(1): 55-63.
  37. Herrington JE. Risk perceptions regarding ticks and Lyme disease: A national survey. *American Journal of Preventive Medicine* 2004; 26: 135-40.
  38. Malouin R, Winch P, Leontsini E, et al. Longitudinal evaluation of an educational intervention for preventing tick bites in an area with endemic Lyme disease in Baltimore County, Maryland. *American Journal of Epidemiology* 2005; 157: 1039-51.
  39. Phillips CB, Liang MH, Sangha O, et al. Lyme disease and preventive behaviors in residents of Nantucket Island, Massachusetts. *Am J Prev Med* 2001; 20: 219-24.
  40. Hochbaum GM. Public participation in medical screening programs: A socio-psychological study (Public Health Service Publication No. 572). Washington, DC: Government Printing Office, 1958.
  41. Jackson LE, Hilborn ED, Thomas JC. Towards landscape design guidelines for reducing Lyme disease risk. *Int J Epi* 2006; 35(2): 315-22.
  42. Radeloff VC, Hammer RB, Stewart SI. Rural and suburban sprawl in the U. S. Midwest from 1940 to 2000 and its relation to forest fragmentation. *Conserv Biol* 2005; 19(3): 793-805.
  43. McMichael AJ. Environmental and social influences on emerging infectious diseases: Past, present, and future. *Philos T Roy Soc B* 2004; 359: 1049-58.
  44. Cromley EK, Cartter ML, Mrozinski RD, Ertel SH. Residential setting as a risk factor for Lyme disease in a hyperendemic region. *Am J Epi* 1998; 147: 472-7.
  45. Farmer P. Social inequalities and emerging infectious diseases. *Emerg Infect Dis* 1996; 2(4): 259-269.
  46. Killilea ME, Swee A, Lane RS, Briggs CJ, Ostfeld RS. Spatial dynamics of Lyme disease: A review. *Ecohealth* 2008; 5(2): 167-95.
  47. Cromley RG, Cromley EK. Choropleth map legend design for visualizing community health disparities. *International J Health Geo* 2009; 8(52): 1-11. doi: 10.1186/1476-072X-8-52
  48. Jackson LE, Levine JF, Hilborn ED. A comparison of analysis units for associating Lyme disease with forest edge habitat. *Community Ecol* 2006; 7: 189-97.
  49. Kilpatrick AM, Randolph SE. Drivers, dynamics, and control of emerging vector-borne zoonotic diseases. *Lancet* 2012; 380(9857): 1946-1955. doi: 10.1016/S0140-6736(12)61151-9
  50. Piesman J, Gage KL. Bacterial and rickettsial diseases. In: Eldridge BF, Edman JD, editors. *Medical entomology*. Netherlands: Springer Science and Business Media B. V., 2000: 377-413.
  51. Hewitt K. *Regions of risk: A geographical introduction to disasters*. Essex, UK: Longman, 1997.
  52. Platt R. *Disasters and democracy: The politics of extreme natural events*. Washington, DC: Island Press, 1999.
  53. Pulido L. Rethinking environmental racism: White privilege and urban development in southern California. *Ann Assoc Am Geo* 2000; 90: 12-40.

54. Hazards and Vulnerability Research Institute. Social Vulnerability Index (SoVI) for the United States: 2006-2010 data. Retrieved from: [http://webra.cas.sc.edu/hvri/products/sovi2010\\_data.aspx](http://webra.cas.sc.edu/hvri/products/sovi2010_data.aspx) [Accessed January 24, 2016].
55. Cutter SL, Solecki WD. The national pattern of airborne toxic releases. *Professional Geo* 1989; 41: 149-61.
56. Agency for Toxic Substances and Disease Registry. The social vulnerability index. Retrieved from: <http://svi.cdc.gov> [Accessed January 28, 2016].
57. Chakraborty J, Tobin GA, Montz BE. Population evacuation: Assessing spatial variability in geophysical risk and social vulnerability to natural hazards. *Nat Hazard Rev* 2005; 6(1): 23-33.
58. Cutter SL, Emrich CT. Moral hazard, social catastrophe: The changing face of vulnerability along the hurricane coasts. *Annal Am Acad Politic Soc Sci* 2006; 604(102): 102-12.
59. Flanagan BE, Gregory EW, Hallisey EJ, Heitgerd JL, Lewis B. A social vulnerability index for disaster management. *J Homel Secur Emerg* 2011; 8(1): 1-22. doi: 10.2202/1547-7355.1792
60. Gaiher CJ, Goodrick S, Murphy BE, Poudyal N. An exploratory spatial analysis of social vulnerability and smoke plume dispersion in the U. S. south. *Forests* 2015; 6(5): 1397-1421. doi: 10.3390/f6051397
61. Koks EE, Jongman B, Husby TG, Botzen WJW. Combining hazard, exposure and social vulnerability to provide lessons for flood risk management. *Environ Sci Policy* 2015; 47: 42-52.
62. Maier G, Grundstein A, Jang W, Li C, Naeher LP, Shapherd M. Assessing the performance of a vulnerability index during oppressive heat across Georgia, United States. *Weather Clim Soc* 2014; 6: 253-63. doi: 10.1175/WCAS-D-13-00037.1
63. Myers CA, Slack T, Singelmann J. Social vulnerability and migration in the wake of disaster: The case of Hurricanes Katrina and Rita. *Pop Environ* 2008; 29(6): 271-91.
64. Rygel L, O'Sullivan D, Yarnal B. A method for constructing a social vulnerability index: An application to hurricane storm surges in a developed country. *Mitigation Adapt Strat Global Change* 2006; 11: 741-64. doi: 10.1007/s11027-006-0265-6
65. Kugeler KJ, Farley GM, Forrester JD, Mead P. Geographic distribution and expansion of human Lyme disease, United States. *Emerg Infect Dis* 2015; 21(8): 1455-7.
66. United States Census Bureau. County intercensal estimates. Retrieved from: <https://www.census.gov/popest/data/intercensal/county/county2010.html> [Accessed December 14, 2016].
67. United States Census Bureau. SVI data and tools download. Retrieved from: <https://svi.cdc.gov/SVIDataToolsDownload.html> [Accessed December 16, 2016].
68. United States Census Bureau. Census regions and divisions of the United States. Retrieved from: [http://www2.census.gov/geo/pdfs/maps-data/maps/reference/us\\_regdiv.pdf](http://www2.census.gov/geo/pdfs/maps-data/maps/reference/us_regdiv.pdf) [Accessed February 3, 2016].
69. Waller LA, Wilson ML, Marshall SL. Spatio-temporal patterns in county-level incidence and reporting of Lyme disease in the northeastern United States, 1990-2000. *Env Ecol Stat* 2007; 14: 83-100.
70. Lopez-Zetina J, Lee H, Friis R. The link between obesity and the built environment: Evidence from an ecological analysis of obesity and vehicle miles of travel in California. *Health Place* 2006; 12(4): 656-64.
71. Gaddum JH. Lognormal distributions. *Nature* 1945; 159: 463-66.
72. Freedman DA. Bootstrapping regression models. *Ann Stat* 1981; 9(6): 1218-28.
73. QGIS Development Team. QGIS geographic information system: Open source geospatial foundation. Retrieved from: [www.qgis.org](http://www.qgis.org) [Accessed December 1, 2016].
74. Hoy M, Livernois J, McKenna C, Rees R, Stengos T. *Mathematics for economics* (2nd ed.). Cambridge, MA: The MIT Press; 2001.
75. Abelson RP. A variance explanation paradox: When a little is a lot. *Psych Bull* 1985; 97(1): 129-33.
76. World Health Organization. International travel and health: Lyme Borreliosis. Retrieved from: [www.who.int/ith/diseases/lyme/en/](http://www.who.int/ith/diseases/lyme/en/) [Accessed December 16, 2016].
77. Poland GA. Vaccines against Lyme disease: What happened and what lessons can we learn? *Clin Infect Dis* 2011; 52(suppl 3): 253-8.
78. United States Senate. Lyme and tick-borne disease prevention, education, and research act of 2015. Retrieved from: [www.congress.gov/bill/114th-congress/senate-bill/1503/](http://www.congress.gov/bill/114th-congress/senate-bill/1503/) [Accessed December 16, 2016].

