Quantifying the relationship between human Lyme disease and *Borrelia burgdorferi* exposure in domestic dogs

Yan Liu,1 Shila K. Nordone,2 Michael J. Yabsley,3,4 Robert B. Lund,5 Christopher S. McMahan,5 Jenna R. Gettings3,5

1School of Community Health Sciences, University of Nevada, Reno, NV; 2Department of Molecular and Biomedical Sciences, Comparative Medicine Institute, North Carolina State University, College of Veterinary Medicine, Raleigh, NC; 3Southeastern Cooperative Wildlife Disease Study, Department of Population Health, College of Veterinary Medicine, University of Georgia, Athens, GA; 4Warnell School of Forestry and Natural Resources, University of Georgia, Athens, GA; 5School of Mathematical and Statistical Sciences, Clemson University, Clemson, SC, USA

Abstract

Lyme disease (LD) is the most common vector-borne disease in the United States. Early confirmatory diagnosis remains a challenge, while the disease can be debilitating if left untreated. Further, the decision to test is complicated by under-reporting, low positive predictive values of testing in non-endemic areas and travel, which together exacerbate the difficulty in identification of newly endemic areas or areas of emerging concern. Spatio-temporal analyses at the national scale are critical to establishing a baseline human LD risk assessment tool that would allow for the detection of changes in these areas. A well-established surrogate for human LD incidence is canine LD seroprevalence, making it a strong candidate covariate for use in such analyses. In this paper, Bayesian statistical methods were used to fit a spatio-temporal spline regression model to estimate the relationship between human LD incidence and canine seroprevalence, treating the latter as an explanatory covariate. A strong non-linear monotonically increasing association was found. That is, this analysis suggests that mean incidence in humans increases with canine seroprevalence until the seroprevalence in dogs reaches approximately 30%. This finding reinforces the use of canines as sentinels for human LD risk, especially with respect to identifying geographic areas of concern for potential human exposure.

Introduction

Lyme disease (LD), caused by the spirochete *Borrelia burgdorferi*, remains the most commonly reported vector-borne disease in the contiguous United States (US) (Mead, 2015; Rosenberg et al., 2018). *Borrelia burgdorferi* is transmitted among vertebrate host species by ticks of the genus *Ixodes*, with both humans and dogs as incidental hosts (Lane et al., 1991). Although neither humans nor dogs are significant in the maintenance of the pathogen and the subsequent transmission to other ticks (Mather et al., 1994), both species can suffer clinical consequences as a result of infection. Early infection in humans often presents with erythema migrans (red rash) and flu-like symptoms. As the infection becomes disseminated in the body, clinical signs associated with the nervous system, heart and joints may manifest (Steere, 2001). With the exception of erythema migrans, the clinical presentation in dogs is similar involving fever, lethargy, arthritis, joint swelling and lymphadenopathy (Little et al., 2010).
Cases affecting the heart, kidneys and the nervous system have been reported, but these manifestations are not commonly reported (Littman et al., 2006; Little et al., 2010; Detmer et al., 2016). The potential severity and chronicity of LD, in humans motivates the need for an accurate understanding of its incidence and distribution, with the main objective being to facilitate early diagnosis and allocation of educational resources necessary for disease prevention.

The Centers for Disease Control and Prevention (CDC) in the US recognizes standard 2-tiered testing (STTT) as the gold standard for confirmatory diagnosis of LD recommending its use since 1995. The first-tier STTT utilizes an immunofluorescence assay or enzyme immunoassay that, when positive or equivocal, should be followed by a second-tier immunoblot (IgM and/or IgG) test (Schriefer, 2015). Current recommendations to initiate STTT are driven by pretest probability in the clinical setting, with exposure history being a critical factor in deciding to test (Moore et al., 2016). Despite human LD being notifiable in the US, under-reporting of infection has been cited as a concern (Nelson et al., 2015; Tseng et al., 2015; Moore et al., 2016; White et al., 2016), most notably in paediatric patients (White et al., 2016). Under-reporting is common in most passive infectious disease surveillance systems, particularly when only high-risk regions are studied, and data collections are not representative of the entire country. Surveillance of human LD is further complicated by the non-specific nature of some clinical symptoms and the pitfalls of STTT (Moore et al., 2016). The end result has been a reported potential 3- to 12-fold discrepancy ratio between the number of CDC-confirmed cases and the number of infections actually occurring in the population (Hinkley et al., 2014; Nelson et al., 2015). Furthermore, this under-reporting may delay recognition of newly endemic regions.

Domestic dogs can serve as a proxy for the risk of exposure to infected ticks for humans. Although the risk of exposure may vary between humans and dogs, an infected dog has typically been exposed to all the necessary conditions for transmission of the pathogen to an incidental host, including humans; these factors include those associated with the vector such as density of infected nymphs (Pepin et al., 2012; Eisen et al., 2016b), presence of reservoir and maintenance hosts (Ley et al., 1995; Schauber et al., 2005; Kilpatrick et al., 2014), climate (Schuaber et al., 2005; Moore et al., 2014; Watson et al., 2017), landscape and environment (Ley et al., 1995; Pepin et al., 2012; Tuite et al., 2013; Seukep et al., 2015; Walter et al., 2016), demographics and socioeconomics (Tuite et al., 2013; Seukep et al., 2015) and geographical location (Tuite et al., 2013). The use of pathogen exposure in dogs as an indicator of human incidence has been shown to be sound in several small-scale studies (Lindenmayer et al., 1991; Rand et al., 1991; Guerra et al., 2001; Johnson et al., 2004; Duncan et al., 2005; Stone et al., 2005) and more recently at the national level (Mead et al., 2011).

While the aforementioned studies are supportive of the concept that canine seroprevalence can be used to assess risk of human exposure to infected ticks, their design and methodology limit actionable decision-making by human healthcare providers. To the point, analyses in these prior studies focused on either correlation statistics or employed simple linear regression techniques to explore associations. Importantly, previous analyses fail to allow for a thorough examination of the relationship at low canine seroprevalence. In particular, the assumption of a linear association leads to inaccurate conclusions (e.g. over- and under-estimation) and fails to describe the dynamic nature of LD over the past decade in which human cases are rising in numbers and expanding in geographic distribution (Kugeler et al., 2015), most notably in transitional regions bordering endemic areas of disease. Dynamic changes in LD have been associated with changes in the range of I. scapularis (Eisen et al., 2016a); however, annual assessment of vector and disease distribution remain logistically and financially unfeasible. Dogs may be a more sensitive marker of the presence of B. burgdorferi at the leading edge of these changes (Mead et al., 2011) and an understanding how their exposure relates to human cases is therefore of significance. In light of recent observations, this study updates, refines and hones an analysis that was published by Mead and colleagues performed at the state and county level using canine data from 2001-2006 (Mead et al., 2011). Presented here are data from 2012 through 2016, a more contemporary observation of the current state of Lyme disease in the US. In addition, the canine test is used more routinely during this time period compared to the early 2000s, providing a broader sample population in terms of geography and number of dogs tested (over 16 million, compared to the approximately 1 million previously considered).

Data availability has also improved the scope for human LD diagnosis in the last 5 years, in part due to growing awareness of disease incidence and testing (Rosenberg et al., 2018). Recent spatial analysis revealed clusters of human cases near the state borders in Maryland, Pennsylvania, Virginia and West Virginia, regions that border established endemic regions of LD (Hendricks et al., 2017). These researchers highlight the need for exploratory surveillance approaches to mitigate the extent at which state level reporting affects accurate estimation of LD progression. We similarly reasoned that exploratory spatial data analyses are necessary for public health professionals on a national level to establish baseline risk assessment and initiate surveillance to monitor changes in LD occurrence. To that end, we estimated the relationship between canine B. burgdorferi seroprevalence (Watson et al., 2017) and human LD incidence (Nelson et al., 2014; NNDSS, 2017) using updated information and a spatio-temporal model that can quantify both linear and non-linear associations, thus avoiding erroneous assumptions regarding the form of the relationship. By demonstrating the strength of the association, particularly in areas of low canine seroprevalence, we can open the discussion on the utility of canine data as an explanatory covariate in spatial analyses that could be used to guide pretest probability and evidence-based testing decisions and identify areas of concern to facilitate public health awareness and educational campaigns.

Materials and Methods

Data structure

The observed canine data consist of test results from 16,595,623 B. burgdorferi modified enzyme-linked immunosorbent assays (ELISAs) performed from 2012 to 2016 in the contiguous US based on the following assays: SNAP®4Dx® and SNAP®4Dx® Plus (IDEXX Laboratories, Inc., Westbrook, ME, USA) available from the Companion Animal Parasite Council website (CAPC, 2017). SNAP®4Dx® Plus is a combination test that detects antibodies against B. burgdorferi, Anaplasma spp., Ehrlichia spp., and antigen from Dirofilaria immitis. The county in which the veterinary clinic is located is reported with the test result. These data are collated through the use of software that con-
nects clinics to IDEXX, and all test results (negative or positive) captured by the network are added to the database on a monthly basis. The ELISA is a qualitative test that detects antibodies that are seroreactive to the C6 peptide which is based on a region of the outer membrane protein (VlsE) of \textit{B. burgdorferi} (Stillman \textit{et al.}, 2014) and has a reported sensitivity and specificity of 96.7\% and 98.8\%, respectively. Similar to human infection, detection of seroconversion to C6 antigen in dogs is indicative of disseminated infection (Bacon \textit{et al.}, 2003; Nyman \textit{et al.}, 2006; Levy \textit{et al.}, 2008; Wagner \textit{et al.}, 2012; Embers \textit{et al.}, 2016). Seroconversion can occur as early as 3-4 weeks post-infection if the bacterial burden is high (Wagner \textit{et al.}, 2012); however, in the data used in this study, the time since infection was not known and cannot be determined from the test results. After treatment, antibodies will decline over time (Levy \textit{et al.}, 2008), but persist for months (or years) in untreated animals. It is also important to note that the C6 ELISA does not detect antibodies to current vaccines (Ting Liang \textit{et al.}, 2000), so vaccinated dogs will not test positive with the exception of prior infections or vaccine failure and subsequent disseminated infection.

The predominate population represented by the data presented here was that of owned dogs brought to a veterinarian for care. The reason for testing was not provided and may include wellness screening, suspicion of disease (e.g., heartworm disease, Lyme disease, anaplasmosis, ehrlichiosis), or monitoring response to therapy. Other demographic details, such as age, sex, or breed, were also not available. Travel and testing history were also not available, so location of exposure is not known. At the population level, it is reasonable to make the assumption that most dogs were exposed within the county and during the year of testing. Some observations may, however, have been misclassified by county or year. It is expected that there are few of these misclassifications with most representing animals that could have been exposed in a neighbouring county and/or previous year. Thus, given the strong spatio-temporal dependence, it is expected that these misclassifications will not unduly influence the analysis.

Overall, 1,062,524 tests were positive, yielding an empirical seroprevalence of 6.40\%. For the purposes of fitting the model and matching the spatio-temporal granularity of the CDC data, the canine test data were aggregated by county and by year. Figure 1 depicts the overall empirical seroprevalence over the 2012-2016 time period for each county. The second data set consisted of confirmed and probable human cases of LD reported by the CDC from 2012 to 2016 at the county level (CDC, 2017). All tests during this period use the same case definition, which is available on the CDC website (NNDSS, 2017). The overall incidence per 100,000 humans by county for the 2012-2016 time period are depicted in Figure 2.

**Statistical models**

Our statistical analysis proceeded in two steps. First, a Bayesian spatio-temporal model was fitted to the canine test data to obtain seroprevalence estimates at each county throughout the contiguous US at all years of interest. The second step analyzed the human incidence data via a different model using the seroprevalence estimates from the first step as an explanatory factor. The overarching goal of this analysis was to accurately quantify the association between human incidence and canine exposure preva-
lence. In the first step, we used the Bayesian hierarchical spatio-temporal model, the salient features of which have been described by Banerjee et al. (2014) and Watson et al. (2017). Let $Y_{st}$ and $n_{st}$ denote the number of positive and total number of canine tests in county $s$ during year $t$, respectively, for counties $s \in \{1, \ldots, S\}$ and years $t \in \{1, \ldots, T\}$. Our step one model was a space-time Poisson regression:

$$Y_{st} | n_{st}, p_{st} \sim \text{Poisson} \{n_{st} p_{st}\},$$

Eq. 1

$$\log \{p_{st}\} = \beta_0 + \xi_{st},$$

Eq. 2

where $\beta_0$ is an intercept parameter and $\xi_{st}$ a spatio-temporal random effect for counties at time $t$. Here $p_{st}$ is the seroprevalence of $B. burgdorferi$ antibodies in dogs in counties at time $t$. The spatio-temporal dependence in the data are described through a multivariate autoregressive model obeying:

$$\xi_1 = \phi_1;$$

Eq. 3

$$\xi_t = \phi \xi_{t-1} + \phi_t, \text{ for } t \in \{2, \ldots, T\};$$

Eq. 4

$$\phi_t \sim \text{CAR} \left( \tau^2; \rho \right), \text{ for } t \in \{1, \ldots, T\}$$

Eq. 5

where $\xi_1 = (\xi_{11}, \ldots, \xi_{S1})'$ and $\phi_t = (\phi_{1t}, \ldots, \phi_{St})'$ are random vectors ($'$ denotes matrix transpose). In this model, the correlation parameter $\phi$ lies between $(-1,1)$ and controls the degree of temporal correlation in the model. Values of this parameter near 1(-1) indicate strong positive (negative) temporal correlation and values near zero indicate weaker temporal correlation. Similarly, the parameter $\rho$ controls the spatial correlation and is constrained to lie between (0,1). Values of this parameter close to unity indicate strong positive spatial dependence, while values close to zero indicate weak spatial correlation. Eq. 5 specifies that $\phi_t$ are independent and identically distributed (iid) random vectors, whose distribution follows a conditional autoregressive (CAR) model. See Besag (1974), Banerjee et al. (2014) and Watson et al. (2017) for details and the statistical properties of CAR schemes. To complete the Bayesian model formulation, the following prior distributions were used:

$$\beta_0 \sim \text{N} \left(0,1000\right),$$

Eq. 6

$$\phi \sim \text{Uniform} \left(-1,1\right);$$

Eq. 7

$$\rho \sim \text{Uniform} \left(0,1\right);$$

Eq. 8

$$\tau^2 \sim \text{Gamma} \left(0.5,0.05\right).$$

Eq. 9

These specifications put a diffuse prior on $\beta_0$; i.e. the prior exerts little influence on the posterior distribution and inferences are based primarily on the data. Uninformative (uniform) prior distributions were used for $\phi$ and $\rho$ for the same reasons. The prior

Figure 2. Average empirical human Lyme disease incidence per 100,000 population by county for 2012-2016. County-level data are publicly available from the Centers for Disease Control and Prevention. The overall incidence for the five-year period is displayed for each county reporting data.
for \( r^2 \) was chosen to be uninformative with statistical conjugacy (Banerjee et al., 2014) in mind, which facilitates computation.

The unknown parameters in the model were \( \beta_0, \theta, \tau, \xi_t, \) for \( s \in \{1, \ldots, S\} \) and \( t \in \{1, \ldots, T\} \). A posterior sampling algorithm was developed, in the usual statistical fashion, to sample all model parameters and random effects from their posterior distributions. This Markov chain Monte Carlo (MCMC) sampling algorithm uses a combination of Gibbs and Metropolis-Hastings steps. To complete model fitting, \( V_j \) for counties not reporting test results were treated as latent variables and were sampled along with the other model parameters. The posterior sampling algorithm was implemented using R and C++. For more information on Bayesian models and MCMC methods, see Gelman et al., 2014. In what follows, an estimate, denoted by \( \hat{p}_{st} \), of \( p_{st} \) was obtained as the sample mean from \( G \) posterior samples \( \hat{p}_{st}^{(g)} \) for \( g \in \{1, \ldots, G\} \), where \( \hat{p}_{st}^{(g)} = \exp\{\hat{g}_t^{(g)} + \hat{\phi}^{(g)}\} \) with \( \hat{g}_t^{(g)} \) and \( \hat{\phi}^{(g)} \) denoting the \( g \)th posterior of \( \beta_s \) and \( \xi_t \), respectively. The Bayesian spatio-temporal model for the second step of this analysis is similar to the first model, but has several key differences. Let \( \xi_t \) and \( m_s \) denote the number of CDC-confirmed human Lyme cases and the population size of county \( s \) during year \( t \), respectively, for counties \( s \in \{1, \ldots, S\} \) and years \( t \in \{1, \ldots, T\} \). Our model for this step was another Poisson regression:

\[
X_{st} | m_s, r_s \sim \text{Poisson} \{m_s r_s\}, \quad \text{Eq. 10}
\]

\[
\log(r_s) = \gamma_0 + \beta_s + \delta_s, \quad \text{Eq. 11}
\]

where \( \gamma_0 \) is an intercept parameter, \( \beta_s \) an unknown function and \( \delta_s \) a spatio-temporal random effect. In the above, \( r_s \) is the risk of a human Lyme disease case in county \( s \) at time \( t \). Specifications for the intercept and spatio-temporal random effects are as described for the step one model. The key quantity here was the function \( g(p) \), as this quantifies the association between canine and human risk. One may posit that \( g(p) \) is non-decreasing in \( p \in [0,1] \); i.e. the mean incidence in humans increases with canine seroprevalence. The unknown function in Eq. 11 represents an infinite dimensional parameter. To reduce the dimensions of the problem, our approach used regression splines to approximate \( g(\cdot) \), a classic statistical practice (Ruppert et al., 2003). In particular, the monotone splines of Ramsay (1988) were utilized:

\[
g(p) = \sum_{l=1}^{L} \gamma_l l_s(p), \quad 0 \leq p \leq 1, \quad \text{Eq. 12}
\]

where the \( l_s(\cdot) \) are integrated spline basis functions and the \( \gamma_l \) are spline coefficients. Once a user specifies the degree of the polynomials and the knot set, the \( L \) basis functions are fully determined; see Ramsay (1988) for further details. The degree controls function smoothness, with adequate smoothness typically being attained with degree 2 or 3. The knot set controls the flexibility of the approximation, with more knots giving more flexibility to adapt to the data. This makes knot selection slightly more tenuous as too few (many) knots may lead to under- (over-) fitting problems. A common statistical approach used to circumvent this problem involves first specifying a large collection of knots and then regularizing the spline coefficients; i.e., the non-useful spline coefficients are drawn toward zero during the estimation procedure. Proceeding in this fashion guards against both over- and under-fitting. To complete the model formulation, the priors in Eq. 6 need to be specified. To perform regularization, the generalized double Pareto shrinkage prior in Armagan et al. (2013) was chosen for the \( \gamma_l \)’s. The degree of the regression splines was set to be 3 and 5 equally spaced knots were specified between the minimum and maximum of the \( \hat{p}_{st}^{(g)} \)’s. Model fitting and parameter estimation was conducted in the same fashion as in the step one model. The model estimates are displayed as maps for ease of interpretation after Kriging (implemented in ArcGIS using the default parameters). Kriging is a spatial interpolation method, in which the values are modelled by a Gaussian process; for further details see Deutsch and Journel (1992). To demonstrate that the model and methodology were robust to spurious outcomes in the canine data (likely attributable to travel-related exposure) observed in non-endemic regions, the model was also fitted using a subset of the described data. This subset included only states classified as high incidence by the CDC (CDC, 2017) and those immediately surrounding: Maine, New Hampshire, Vermont, Massachusetts, Connecticut, Rhode Island, New York, New Jersey, Pennsylvania, Delaware, Maryland, Virginia, West Virginia, North Carolina, Tennessee, Kentucky, Ohio, Indiana, Michigan, Illinois, Iowa, Wisconsin, Minnesota, North Dakota, South Dakota and Missouri.

**Table 1. Estimates of parameters.**

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Estimate (posterior mean)</th>
<th>95% Credible Interval</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \beta_0 )</td>
<td>-12.506</td>
<td>(-12.594, -12.411)</td>
</tr>
<tr>
<td>( \phi )</td>
<td>0.877</td>
<td>(0.852, 0.904)</td>
</tr>
<tr>
<td>( \rho )</td>
<td>0.999</td>
<td>(0.999, 1.000)</td>
</tr>
<tr>
<td>( \tau^2 )</td>
<td>0.814</td>
<td>(0.762, 0.872)</td>
</tr>
</tbody>
</table>
non-endemic regions. Table 1 provides point estimates and 95% credible intervals for the other model parameters.

County level estimates of the number of cases per 100,000 humans can be obtained from the model for each considered year. Figure 5 shows these estimates as the average of the number of cases per 100,000 humans after Kriging. From this map, the states that contain mostly high-incidence areas (10 cases or more per 100,000 humans) were: Minnesota, Wisconsin, Maine, New Hampshire, Vermont, Massachusetts, New York, Connecticut, Rhode Island, New Jersey, Pennsylvania, Maryland, Delaware and Virginia. Bordering states that contain some areas of high-incidence included the Upper Peninsula of Michigan, Iowa, Illinois, West Virginia and North Carolina, while North Dakota and Ohio had the leading edge of high-incidence on their eastern borders. The remainder of the country was found to have low incidence (<10 cases per 100,000 humans) but areas with slightly increased incidence (>1 cases per 100,000 humans) were seen on the leading edges of the endemic regions: the north-western and south-western corners of Indiana and Michigan, respectively, and along the eastern Lake Michigan shore, most of North Carolina and along the Atlantic coast into South Carolina, Northern California and Southern Oregon. Smaller foci were observed in Florida, Southern Kentucky and at the border of Kentucky and Indiana.

Discussion

This paper quantifies the association between canine seroprevalence and human LD incidence within the contiguous US at the county level. The relationship appears well described by a monotone increasing function (denoted by \( g \) previously). A spatio-temporal spline regression model was used to avoid making assumptions about the form of the association. That is, splines are very flexible functions which allow the model to fit the data at all levels of the explanatory covariate, capably illuminating any underlying relationship (Howe et al., 2011). Importantly, regularization was used to prevent the spline model from over-fitting, while spatio-temporal random effects were utilized to account for unexplained spatio-temporal autocorrelation. The estimated association displayed in Figure 4, is our primary result. This analysis demonstrates that mean human incidence monotonically increases as dog seroprevalence increases from 0% to 30%. Notable in this analysis is the narrow credible intervals at low seroprevalence, indicating low variability in the model estimates and a strong association between seroprevalence and human incidence. If dogs were unreliable sentinel for human LD, this interval would be wider, and the association would be interpreted as weaker. Above 30% canine seroprevalence, mean human incidence seems to plateau. The reason for this plateau cannot be elucidated from this study alone and, in part, may be caused by geospatial limitations: there are very few regions of the US where canine seroprevalence is >30%, and with limited data points it is difficult to fully describe the association with human incidence at higher seroprevalence. Beyond these limitations, differential behaviours in hyperendemic counties compared to other regions of the country exist that may impact data acquisition, including but not limited to preventative care practices in both humans and dogs. Estimates derived from Eq. 11 are displayed in Figure 5, which accurately describes incidence while adjusting for the spatial and

![Figure 3. Model-estimated canine seroprevalence, averaged for 2012-2016.](image-url)
temporal random effects unique to each county and demonstrates the ability of seroprevalence to be used as a covariate for spatial models of human incidence. The difference between the model-based estimates and the observed cases of human Lyme disease can be seen in Figure S1 (in the Appendix). The finding that the mean human incidence increases with canine seroprevalence and further that the association is strong at low seroprevalence levels suggest that future research could be aimed at using canine data to monitor areas of transition between endemic and non-endemic regions. This and a previous study by our group (Watson et al., 2017) demonstrate that the widespread availability and timeliness of canine data provides a consistent and spatially uniform surveillance tool in comparison to human data, and therefore enhances the ability to detect areas where *B. burgdorferi* may be diagnosed (either due to travel-acquired infections or due to an area of new transmission of the pathogen). If canine seroprevalence continues to increase, these data would suggest a need to examine the ticks or human case data to determine if endemic transmission was occurring. Future research should focus on the use of a combination of canine data with human case reports and entomological data in defining emergent endemic counties and thus guide pretest probability. For this, a stronger understanding of which specific risk factors influence the association at the county level is needed (Eisen et al., 2003).

Our model does not predict areas of temporal emergence and
the results should not be interpreted as such. Time trends were not included in the model because of the short five-year study period. In light of this limitation, it is currently recommended that canine data be examined annually. Limitations do exist in the data that preclude simple explanations of the association. These include movement of dogs (Millen et al., 2013) that does not follow that of humans (i.e. shelter and rescue dogs); non-human biting *B. burgdorferi* vectors such as *L. affinis* that may bite dogs; cross-reactivity with other borrelial organisms; and false positive test results. Importantly, our model mitigates these factors, as can be seen by comparing the estimated human incidence map (Figure 5) to the observed incidence (Figure 2). However, not all travel-related positive tests should be disregarded. For example, despite Florida being a non-endemic state, many positive tests are reported, and it is easy to conjecture that these are a result of travel between Florida and endemic regions. Since pets do not travel alone, if several canine cases are present within a county, it suggests that the population of that county has some behaviour (travel, camping, hiking) that exposes pets, and therefore likely their owners, to the risk of LD. Using rickettsial disease as an example of a vector-borne disease in which illness in dogs can precede that of their owners (Elchos and Goddard, 2003), monitoring of canine LD prevalence at the county level (CAPC, 2017) and enhanced communication between veterinarians and human health care providers would be optimal when LD is suspected.

The validation of the strength of the association between canine and human LD opens the door to focused studies in areas of concern where the presence of *B. burgdorferi* in ticks is difficult to ascertain and sampling humans is expensive and time consuming. Future considerations will include models that describe the emergence of risk in non-endemic areas. Enhanced sampling of dogs, both owned and non-owned, in combination with the emergence of novel human LD diagnostics (Wormser et al., 2013; Magni et al., 2015; Molins et al., 2015; Embers et al., 2016) that improve upon STTT, have the potential to drive further improvements of national-level risk assessment tools for public health promotion. Data could also be used to guide physicians toward evidence-based use of protective vaccines as they emerge in the marketplace (Izac et al., 2017), with an emphasis on high-risk individuals (children 5-9 and adults 45-59 years of age (CDC, 2017)) and/or potential travel-related exposure. This would ultimately contribute to reducing the financial burden placed on the US health care system by LD (Adrion et al., 2015).

**Conclusions**

Lyme disease can be difficult to diagnose, and chronic manifestations may be severe and debilitating, but it is preventable with appropriate awareness and protection. The model presented here quantifies the association between canine seroprevalence and human LD incidence. Given the availability and timeliness of canine seroprevalence data, we describe a link to human LD incidence that may ultimately serve as an annual early-warning system in transitional and non-endemic regions facilitating awareness and aid in early and accurate diagnosis.

**References**


Hendricks B, Mark-Carew M, 2017. Using exploratory data anal-


