



Original Investigation | Infectious Diseases

Current and Future Spatiotemporal Patterns of Lyme Disease Reporting in the Northeastern United States

Donal Bisanzio, DVM, PhD; Maria P. Fernández, PhD; Elisa Martello, DVM, PhD; Richard Reithinger, PhD; Maria A. Diuk-Wasser, PhD

Abstract

IMPORTANCE The incidence and geographic range of Lyme disease continues to increase in the United States because of the expansion of *Ixodes scapularis*, the species of tick that is the main Lyme disease vector. Currently, no dynamic model for the disease spread exists that integrates information of both acarological and human case surveillance data.

OBJECTIVE To characterize the spatiotemporal spread of Lyme disease in humans among counties in US endemic regions.

DESIGN, SETTING, AND PARTICIPANTS This cross-sectional study modeled the spread of Lyme disease county-level case reporting, accounting for county-level demographic factors, environmental factors associated with tick presence and human exposure, and the spatiotemporal association between counties. The analyses were conducted between January and August 2019. The setting was 1405 counties in the following regions of the United States: West North Central, East North Central, New England, Middle Atlantic, and the South. Assessments were based on publicly available Lyme disease case data reported to the US Centers for Disease Control and Prevention (CDC) between January 2000 and December 2017.

MAIN OUTCOMES AND MEASURES Probability of reporting the first case of Lyme disease by county by year.

RESULTS Between 2000 and 2017, a total of 497 569 Lyme disease cases were reported to the CDC in the study area. Reporting a first case of Lyme disease was associated with a county's and county's neighbors' forest coverage, elevation, percentage of population living in the wildland-urban interface, tick presence, county's population size, proportion of neighbors reporting cases, and neighbors' years since first reporting. The model that included these variables showed high predictive power, with a mean area under the receiver operating characteristic curve of 81.1 (95% CI, 68.5-86.2). The model predicted the first reported Lyme disease case a mean (SD) of 5.5 (3.5) years earlier than was reported to the CDC, with a mean spread velocity estimated at 27.4 (95% CI, 13.6-54.4) km per year. Among 162 counties without reported cases, 47 (29.0%) had a high probability of reporting Lyme disease by 2018. The estimated mean time lag between the first reported case in a neighboring county and any county was 7 (95% CI, 3-8) years.

CONCLUSIONS AND RELEVANCE This study's findings suggest that, if updated regularly and expanded geographically, this predictive model could enable states and counties to develop more specific Lyme disease prevention and control plans, including improved sensitization of the general population and medical community.

JAMA Network Open. 2020;3(3):e200319. doi:10.1001/jamanetworkopen.2020.0319

Key Points

Question What are the characteristics of the spatiotemporal spread of Lyme disease in humans across counties in US endemic regions?

Findings Among 1405 counties with 497 569 Lyme disease cases in this cross-sectional study, the probability of a county reporting its first Lyme disease case between 2000 and 2017 was associated with the county's and the county's neighbors' environmental and demographic factors, tick presence, and neighbors' reporting status.

Meaning This study's findings suggest that model results could allow states and counties to develop more specific Lyme disease prevention and control plans and strategies, including optimizing active surveillance of disease cases and ticks to increase early case detection.

+ [Invited Commentary](#)

+ [Supplemental content](#)

Author affiliations and article information are listed at the end of this article.

Open Access. This is an open access article distributed under the terms of the CC-BY License.

Introduction

Lyme disease is a multisystem zoonotic, vector-borne bacterial infection caused in North America by *Borrelia burgdorferi* sensu stricto that is transmitted by the bite of infected *Ixodes scapularis* or *Ixodes pacificus* ticks.^{1,2} *Borrelia burgdorferi* is maintained in nature in an enzootic cycle comprising multiple vertebrate species; humans are incidentally infected when exposed to infected ticks in forested or highly vegetated areas.^{1,3} With approximately 30 000 reported cases and an estimated 300 000 cases occurring annually,⁴ *B burgdorferi* is the most commonly reported vector-borne disease in the United States, representing 62.6% of cumulative reported vector-borne disease cases.⁵ Most Lyme disease cases occur in the northeastern and North Central states,⁶ where the number of endemic and high-incidence counties continues to expand.^{7,8}

Quantifying the spatiotemporal spread in humans of tick-borne diseases maintained in enzootic transmission cycles remains a major challenge. Although the distribution of infected tick-borne disease vectors has been used to map the entomological or acarological risk of Lyme disease,^{9,10} vector-based data sets are generally not available longitudinally (with some exceptions, including New York State¹¹) and thus are not useful to model disease spread. Furthermore, if the outcome of interest is the geographic distribution or the incidence of human disease, vector data only partially correlate with disease cases.^{9,12}

The patterns of Lyme disease expansion at the national level can be inferred using US Centers for Disease Control and Prevention (CDC)-compiled reported cases based on the standardized case definition approved by the Council of State and Territorial Epidemiologists.¹³ Surveillance for Lyme disease in the United States is based on reports submitted by laboratories and health care providers to state and local health departments.¹⁴ As stated by Nelson et al,⁴ "These reports provide valuable insight into the age and sex distribution of patients with [Lyme disease] and the seasonality and geographic distribution of cases, and they enable monitoring of disease trends over time."^{4(p1625)} Unfortunately, because of biases in diagnosis and reporting, as well as variation in surveillance practices by administrative jurisdictions, reported cases capture only a fraction of the overall frequency of Lyme disease within a population.⁴

Although both overreporting and underreporting occur,^{3,15,16} previous studies^{4,17,18} at the state and national level indicate that underreporting of clinically diagnosed Lyme disease is the dominant net outcome. However, Lyme disease underreporting is not uniform. According to Nelson et al,⁴ although 96.3% of Lyme disease cases reported to the CDC between 2005 and 2010 were concentrated in the 15 highest-incidence states, 19.4% of all clinically diagnosed cases investigated occurred outside of those states, indicating that they were missed in CDC reports. Multiple factors likely lead to increased reporting in high-incidence states, including greater public and health care provider awareness. Furthermore, the fact that a history of exposure in a state or county with at least 2 reported human cases is one of the criteria for considering a Lyme disease case as confirmed may intensify this bias and result in increased underreporting in counties or states where Lyme disease is emerging. A skewed emphasis of surveillance efforts on these high-incidence states and counties may then limit the ability to identify areas of Lyme disease emergence, where misdiagnosis or late diagnosis can lead to more severe disease manifestations.¹⁹

Because no standardized nationwide tick surveillance system currently exists in the United States, there is a need to maximize information extraction from data reported by the CDC in an attempt to overcome potential biases and produce actionable public health information. Previous models of reported Lyme disease spatiotemporal patterns have been generated for New York State,^{18,20} the Northeast,^{8,18} and areas undergoing geographic expansion of the disease, such as Michigan,²¹ Virginia,²² Minnesota,²³ Iowa,²⁴ and Pennsylvania.¹⁹ Our study expands the geographic scope of this work to include all of the main endemic areas in the United States identified by Kugeler et al⁷ as high-incidence states, as well as surrounding states. In addition, a novel modeling approach is presented to describe the progression of Lyme disease at a county level by estimating the yearly probability of reporting the first case of Lyme disease in each county in the study area. This approach

goes beyond describing spatiotemporal patterns of cases to focus on identifying emerging areas based on ecological factors (eg, distribution of tick vectors)⁸ and human behavioral factors (eg, travel history and information sharing between neighboring counties). Specifically, a time series of case-reporting data in focal and neighboring counties was constructed between January 2000 and December 2017, as well as environmental and available tick data mechanistically linked to Lyme disease risk, to estimate associations with the probability of reporting the first case of Lyme disease in disease-free counties. The analyses were conducted between January and August 2019. This approach involved a statistical regression model, the results of which were included in a diffusion model to simulate the spatiotemporal dynamics of the probability of reporting the first Lyme disease case between 2000 and 2017. By incorporating covariates associated with the underlying ecological processes contributing to the disease spread of counties and their neighbors, it was expected that our model approach would identify nonreporting counties with high probability to begin reporting Lyme disease cases in the future.

The objectives of this study were (1) to characterize biotic and abiotic factors that altered the diffusion of Lyme disease reporting between 2000 and 2017 and (2) to identify Lyme disease-free counties with high reporting probability. Our findings can help public health authorities focus surveillance efforts in counties identified in the model as being at high risk of Lyme disease emergence but where cases have yet to be reported.

Methods

Study Setting and Data

This cross-sectional study included 31 states in the following regions of the United States: West North Central, East North Central, New England, Middle Atlantic, and the South²⁵ (eFigure 1 in the [Supplement](#)). These regions were chosen because they include counties with a high incidence of Lyme disease,⁷ counties that started reporting cases since 2000, and counties that have never reported a Lyme disease case. Because the study involved the analysis of preexisting, deidentified data, it was exempt from institutional review board approval by RTI International and Columbia University. This study followed the Strengthening the Reporting of Observational Studies in Epidemiology ([STROBE](#)) reporting guideline.²⁶

Lyme Disease Cases

Our study used publicly available Lyme disease case data from the CDC.²⁷ The reported number of cases at the county level corresponds to the cases reported from state and local health officials to the CDC through the National Notifiable Diseases Surveillance System. Cases are reported according to standardized case definitions. Starting in 2008, the case definition (note that the case definition changed again in 2011 and 2017) of Lyme disease includes probable cases (ie, a case clinically diagnosed with laboratory evidence of infection) and confirmed cases (ie, erythema migrans [the so-called bull's-eye rash] with a known exposure or with laboratory evidence of infection or any case with ≥ 1 late manifestation of Lyme disease symptoms and laboratory evidence of infection).¹³ Exposure is defined as having visited a suitable tick habitat in an endemic county (or in an endemic state after the 2017 change in case definition) 30 days before onset of symptoms. A county (or a state after 2017) is considered endemic when at least 2 cases (10 cases for a state after 2017) of Lyme disease have been reported or there is an established population of infected *Ixodes* ticks. The data set used in this study included all Lyme disease cases reported between 2000 and 2017 in our study area. Therefore, the reported cases included in this study followed the case definitions established before 2017. The publicly available data from the CDC are deidentified in compliance with the Health Insurance Portability and Accountability Act Privacy Rule.

Geographic Database

Counties' boundaries and population estimates between 2000 and 2017 were obtained from the US Census Bureau.²⁵ Forest coverage at the county level was calculated using land cover information obtained from the Multi-Resolution Land Characteristics Consortium.²⁸ Total hectares of territory covered by deciduous, evergreen, and mixed forest were calculated for each county. The fraction of the population living at a wildland-urban interface (WUI) was obtained from the SILVIS Lab website.²⁹ NASA Shuttle Radar Topography Mission satellite images were downloaded from the NASA website³⁰ to calculate the mean elevation of each county. Information about the presence of *I. scapularis* vectors in the study area was extracted from work by Eisen et al.³¹ For each county, a dichotomous variable was created in the geographic database that indicated whether a county had an established tick population or not (either occasional tick reporting or no reporting at all) as of 2015. However, those data were only available for 2000 to 2015, and the 2016 and 2017 estimates include vector data fixed at 2015 (meaning that for 2016 and 2017 the same values as for 2015 were used). All geography data were stored in a geographic database created using the software programs QGIS (version 3.4; OSGeo)³² and GRASS (version 7.2; OSGeo).³³

Statistical Analysis

A 2-step modeling approach was performed to (1) identify the relevant explanatory variables of Lyme disease reporting and the magnitude of their associations and (2) use those findings to simulate the spread of Lyme disease across the study area between 2000 and 2017, assuming a diffusionlike process. The first step consisted of a logistic regression based on generalized linear mixed models to evaluate the probability of reporting the first case of Lyme disease between 2000 and 2015, as described in the next sentence below, according to the following characteristics: a county's population size and tick presence³¹; the percentage of the county covered by deciduous forest, evergreen forest, and mixed forest; the fraction of the population living in WUI areas; and the county's mean elevation. For each county, the regression model also accounted for the characteristics of contiguous counties (ie, first-degree neighbors) and contiguous counties to the first-degree neighbors (ie, second-degree neighbors) to allow for potential exposure to infected vectors in out-of-county locations and the spread of infected vectors into the focal county after a diffusionlike process. The state and year were included in the model as random effects to take into account reporting variations between states and years. The regression model was performed using data between 2000 and 2015 (as training data), and Lyme disease cases reported in 2016 and 2017 were used to validate and test the model predictions. Model selection based on the Akaike information criterion was performed to identify the best model.³⁴ The sensitivity and specificity of the model were estimated by calculating the area under the receiver operating characteristic curve (AUC)³⁵ using the model predictions and the cases reported in 2016 and 2017. A complete description of the logistic regression model and model selection approach is included in the eAppendix in the [Supplement](#).

The second step of the modeling approach consisted of a stochastic, spatially explicit diffusion model to predict the spatiotemporal progression of Lyme disease reporting in the study area between 2000 and 2017. This model consisted of a planar network in which each node was a county connected with the contiguous neighboring counties (eFigure 2 in the [Supplement](#)), and the diffusion of Lyme disease reporting among the counties was assessed by the coefficients obtained from the regression model. The state of Lyme disease reporting in 2000 was used as the baseline, and 1000 simulations were run for 2000 to 2017, randomly removing 100 counties in each simulation to incorporate stochasticity into the model and to reduce reporting bias. At 1-year steps, the model predicted the reporting status of each county based on the characteristics of the county and its contiguous counties. A detailed description of the diffusion model is included in the eAppendix in the [Supplement](#). The outcomes of this model were the probability of reporting at least 1 case of Lyme disease in a county by 2018 (the predicted year for the first case being reported) and the velocity of the spread of Lyme disease. The results of the diffusion model were then compared

with the observed case data obtained from the CDC. A detailed description of the method used to calculate the velocity of the spread of Lyme disease is included in the eAppendix in the [Supplement](#).

The modeling approach accounted for the changes in case definition made by the CDC in 2008 and 2011. To avoid mixing cases reported using different case definitions, the modeling code at the beginning of each simulation randomly used the subset of reported cases in the following time windows: 2000 to 2007, 2008 to 2010, and 2011 to 2015.

Results

Between 2000 and 2017, a total of 497 569 Lyme disease cases were reported to the CDC in the 1405 counties included in our analyses in the following regions of the United States: 27 963 in West North Central, 39 774 in East North Central, 130 992 in New England, 247 029 in Middle Atlantic, and 51 811 in the South. The publicly available data do not include sex, age, or other demographic information.

The model selection showed that the best logistic regression model to predict the occurrence of a first case of Lyme disease between 2000 and 2017 (**Figure 1**) was associated with increasing county's tick establishment (any type), decreasing elevation, increasing percentage population living in a WUI, tick presence, increasing population size, increasing proportion of first-degree neighbors reporting Lyme disease cases and tick presence since 2000, increasing year lag between the first reporting by any first-degree neighboring county, and increasing first-degree neighbors' forest coverage (**Table**). The final formula of the best model is reported in the eAppendix in the [Supplement](#). The model that included these variables showed strong classification performance (ie, predictive power), with a mean AUC of 0.81 (95% CI, 0.69-0.86). The performance of the model was substantially reduced when only county factors were considered (AUC, 0.66; 95% CI, 0.55-0.70) or when the model also included second-degree neighbors' characteristics (AUC, 0.76; 95% CI, 0.64-0.80).

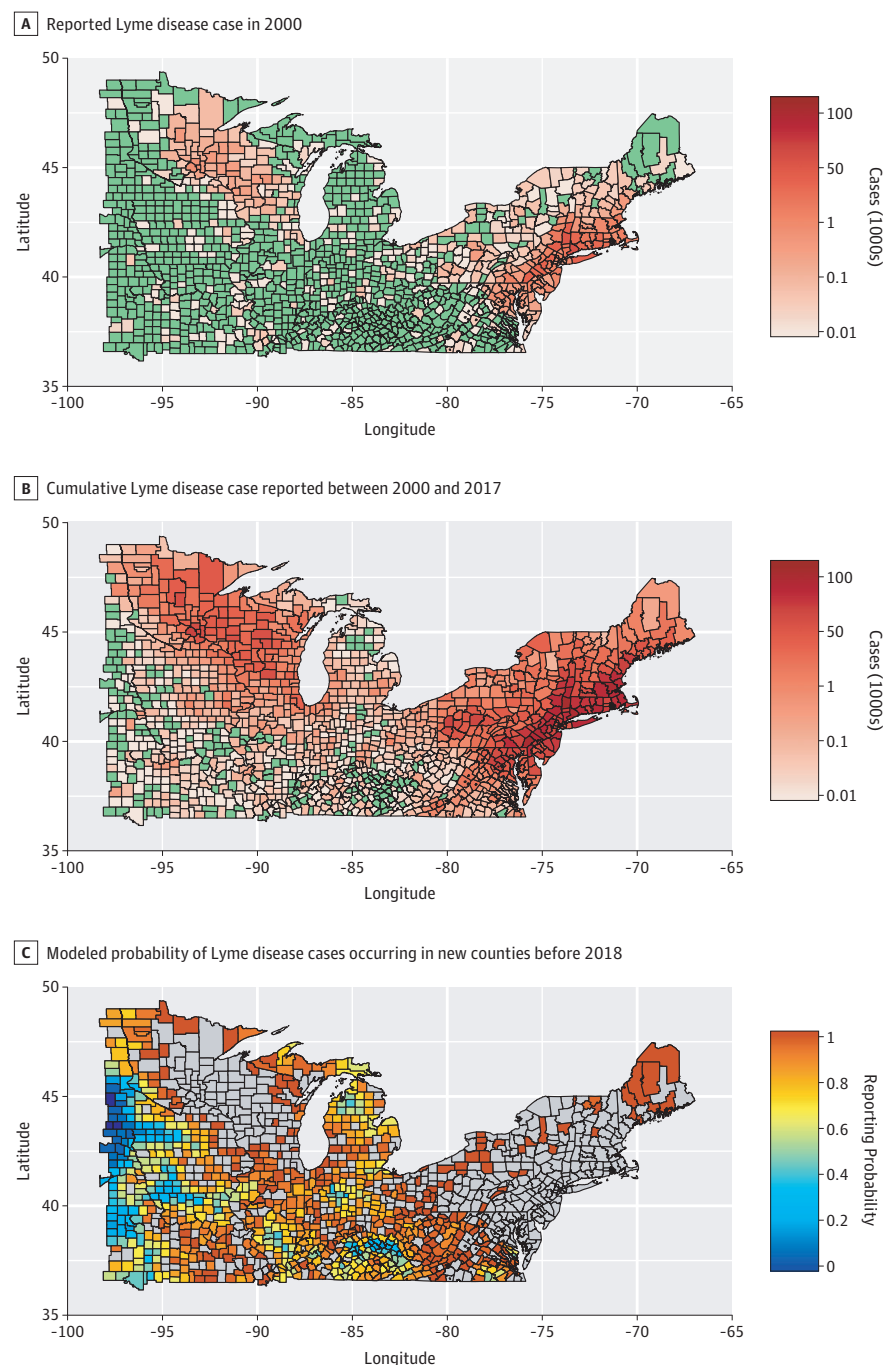
The spread of reported Lyme disease cases predicted by the diffusion model showed high similarity to the observed data reported to the CDC (**Figure 2**). Between 2000 and 2017, Lyme disease mostly spread toward the west and south of the study area, as well as north into Michigan's Lower Peninsula (Figure 2A). Counties that reported cases continuously during the study period had an estimated reporting probability above 0.8 according to the diffusion model. Moreover, counties that first reported a Lyme disease case in 2016 and 2017 (data not included in the logistic regression model) showed a high reporting probability according to the model's predictions (**Figure 3**). However, a mismatch was observed between the model predictions and counties not reporting Lyme disease cases between 2000 and 2017 ($n = 162$): 47 (29.0%) of these counties had a high (>0.8) probability of reporting Lyme disease cases by 2018 (eFigure 3 in the [Supplement](#)). This fraction varied among regions, ranging from 20.7% in West North Central to 39.4% in East North Central to 34.6% in the South.

Timing of first Lyme disease reporting estimated with the model showed a spread pattern similar to that of reported data (Figure 1A and B and Figure 2A and B). Yet, the model estimated a more rapid temporal diffusion, with the first reported Lyme disease case occurring a mean (SD) of 5.5 (3.5) years earlier than the observed data (Figure 2C). The estimated mean time lag between the first reported Lyme disease case in a first-degree neighboring county and any county was 7 (95% CI, 3-8) years (**Figure 4**). However, this lag decreased as the number of reporting first-degree neighboring counties increased (Figure 4). The mean Lyme disease spread velocity (as based on case reporting) was estimated at 27.4 (95% CI, 13.6-54.4) km per year for the entire study area. Diffusion velocity was higher among counties in the South (49.1 km per year; 95% CI, 25.1-70.1 km per year) compared with East North Central (21.1 km per year; 95% CI, 15.3-32.1 km per year) and West North Central (12.2 km per year; 95% CI, 6.4-21.4 km per year) counties.

Discussion

The novel modeling approach used in this study was able to recreate the spatiotemporal pattern of Lyme disease case reporting between 2000 and 2017 by integrating Lyme disease cases reported to the national surveillance system with ecological, environmental, and spatiotemporal covariates. Our findings highlight that forest coverage, elevation, and percentage population living in a WUI are important risk factors for the spread of Lyme disease in the United States. Although such factors were linked to the occurrence of Lyme disease in previous studies reviewed by Killilea et al,³ our study also shows the importance of environmental characteristics of neighboring counties and their

Figure 1. Reported Lyme Disease Cases and Estimated Probability of Case Reporting by 2018



reporting status in predicting the probability of reporting Lyme disease cases. The high contribution of neighboring counties to the reporting probability may be associated with the following 3 processes acting at regional and local levels: (1) diffusion determined mainly by the introduction of the vector and the agent in a suitable habitat; (2) human-driven processes on a local scale, such as tick exposure in neighboring counties; and (3) information sharing across neighboring counties that would increase public and health care provider awareness.

The model estimated that the first report of Lyme disease cases in a county occurred on average 5.5 years earlier than was reported to the CDC, and 29.0% of counties not reporting a case between 2000 and 2017 had an 80% probability of reporting a case by 2018. We hypothesize that this earlier prediction of occurrence by our model could be associated with underreporting, including case misclassification.¹⁸ Such misclassification could include patients not seen with erythema migrans, the cutaneous manifestation of Lyme disease that typically prompts health care providers to inquire about travel history and request laboratory testing.¹⁷ Patients without erythema migrans but with other common noncutaneous manifestations of Lyme disease (eg, chronic arthritis, facial palsy, and cardiac abnormalities) should only undergo laboratory testing for Lyme disease if no other explanation for the symptoms can be found.¹³ However, health care providers working in counties where no Lyme disease cases have been reported before may not consider Lyme disease when performing differential diagnoses.¹⁵

Our model approach included ecological and environmental factors associated with the distribution of tick vectors and human exposure^{8,36} that would predict the occurrence of cases before they are diagnosed and subsequently reported. In emerging areas having limited experience with Lyme disease, underreporting is likely to occur more often because of diagnostic uncertainty or lack of knowledge about reporting requirements.¹⁷ Therefore, the model allowed us to identify potential areas where underreporting might be occurring. Once Lyme disease reporting is established in a county, then our model accurately reflects the spatiotemporal pattern observed in the cases reported to the CDC (ie, counties that first reported a Lyme disease case in 2016 and 2017 showed a high reporting probability according to the model predictions).

Table. Model Ranking Based on Results From Model Selection^a

Characteristic	Model ^b					
	1	2	3	4	5	6
County deciduous forest coverage, hectares	/	/	/	+	+	+
County nondeciduous forest coverage, hectares	/	/	/	+	/	/
County any forest coverage, hectares	+	+	+	/	/	/
NB mean deciduous wood coverage, hectares	+	+	+	/	/	+
County elevation	–	–	–	–	+	+
NB mean nondeciduous forest coverage, hectares	–	+	+	/	/	/
County population size	+	+	+	+	+	+
NB any forest coverage, hectares	+	+	/	/	/	/
NB elevation	/	/	/	+	+	+
No. of reporting NBs	+	+	/	/	/	/
Time lag from first reporting NB	+	+	/	/	+	/
County % population living in WUI	+	/	+	+	/	+
NB mean % population living in WUI	/	+	/	/	+	/
County vector established	+	+	+	+	+	+
No. of NBs with established vector	+	+	+	/	+	/
Frequency of being selected as best model	0.81	0.11	0.05	0.04	0.01	0.01

Abbreviations: NB, first-degree neighboring county; WUI, wildland-urban interface; +, positive association; –, negative association; /, not included in the model.

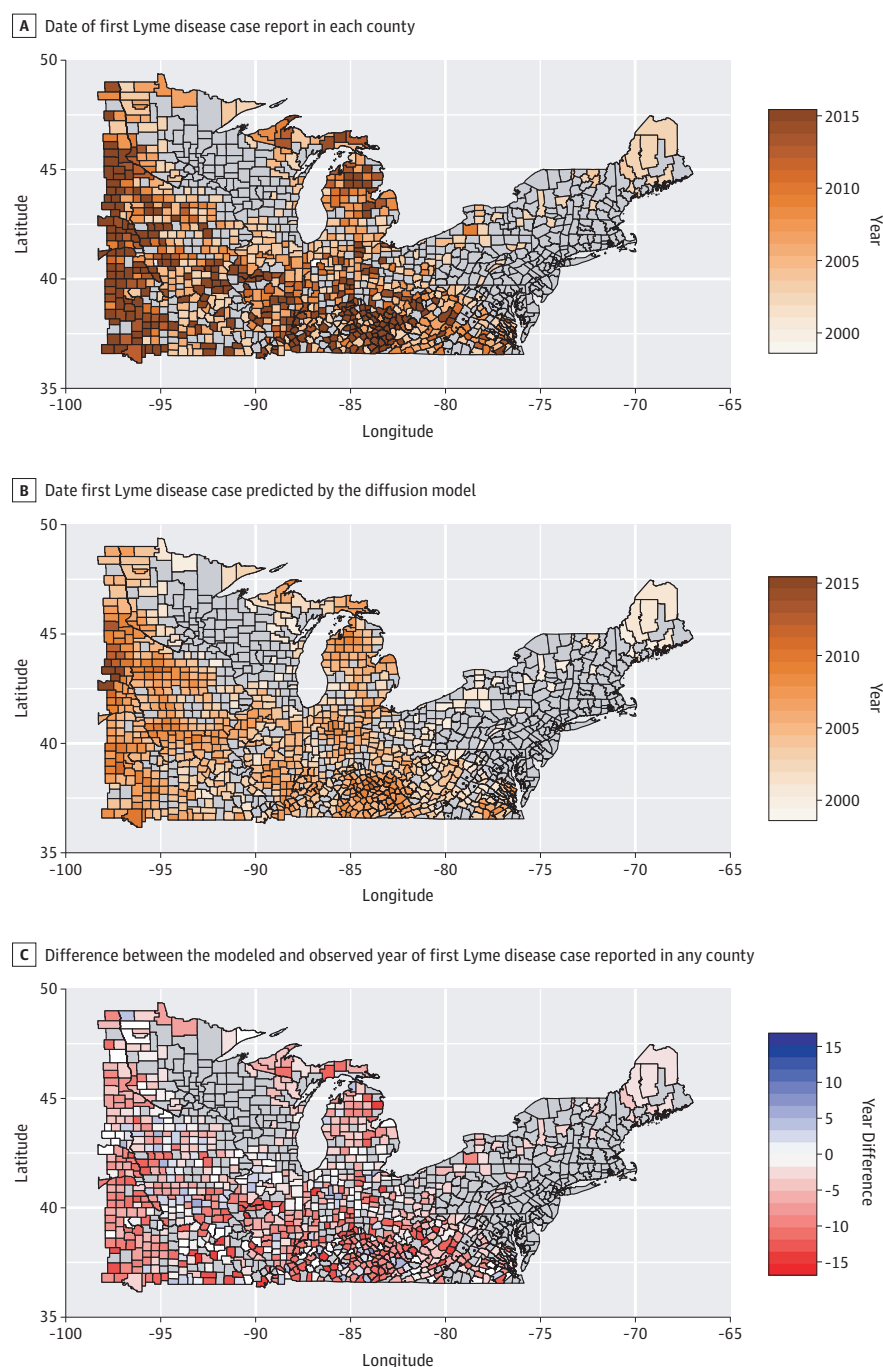
^a Model ranking is based on the model's descending frequency of being flagged as the best model in 1000 repetitions. Each model included the state and year as random effects to account for differences in reporting at the state level, as well as changing Lyme disease case definition between 2000 and 2015. Each model also included the

population size to adjust for the increased probability to record a case of Lyme disease in counties with high populations.

^b A positive association means that with increasing values the greater the probability to report a Lyme case for the first time. A negative association means that with decreasing values the greater the probability to report a Lyme case for the first time.

Our study identified that about one-third of the counties not reporting Lyme disease by 2017 had a high probability of reporting cases by 2018. This finding offers an opportunity for public health officials to anticipate and explore geographic areas where Lyme disease may be occurring or could emerge in the future. This model may be updated yearly as more cases are reported to the CDC, and the results could be used to direct active surveillance efforts to explore underreporting issues, for example.¹⁷ Moreover, combining case reporting with tick surveillance would offer insights into local *B burgdorferi* transmission.³⁷ Our approach could help public health institutions prioritize counties

Figure 2. First Year of Lyme Disease Case Reporting and Modeled Spread of Lyme Disease



Timing of first Lyme disease reporting estimated with the model showed a spread pattern similar to that of data reported to the US Centers for Disease Control and Prevention.

for active tick surveillance or promote passive tick reporting among the general public. As tick data are collected, they may be used to expand the model and iteratively improve the model predictions.

Some differences were observed between regions regarding the diffusion dynamics of Lyme disease reporting, even after accounting for potential reporting biases among states. Although almost all counties located in the Northeast had a high probability of reporting a case between 2000 and 2017, the fraction of counties with high probability of reporting but with no observed cases by 2017 differed among the remaining regions: it was higher in the counties located in the South and the East North Central region compared with the West North Central region. Likewise, the spread velocity estimated by the model was faster in the South and the East North Central region and slower in the West North Central region. These results highlight the importance of the South and the East North Central region as areas of active spread of Lyme disease, at least at the county level, a finding that would benefit from further research to assess underlying explanatory factors.

Limitations

This study has several limitations given the nature of the reported data that were publicly available and used in our analyses. Changes in case definitions between periods and in reporting practices across jurisdictions were accounted for by including the year and the state as covariates, but we cannot rule out additional differences in reporting practices on other levels (ie, county and local health departments). Also, Lyme disease cases are reported in the county where they are diagnosed

Figure 3. Estimated Probabilities of Counties to Report a Lyme Disease Case by 2018 if They Had Reported Lyme Disease Cases for the First Time in 2016 and 2017

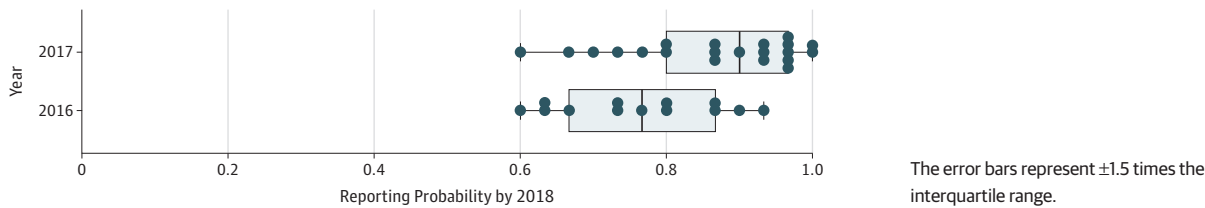
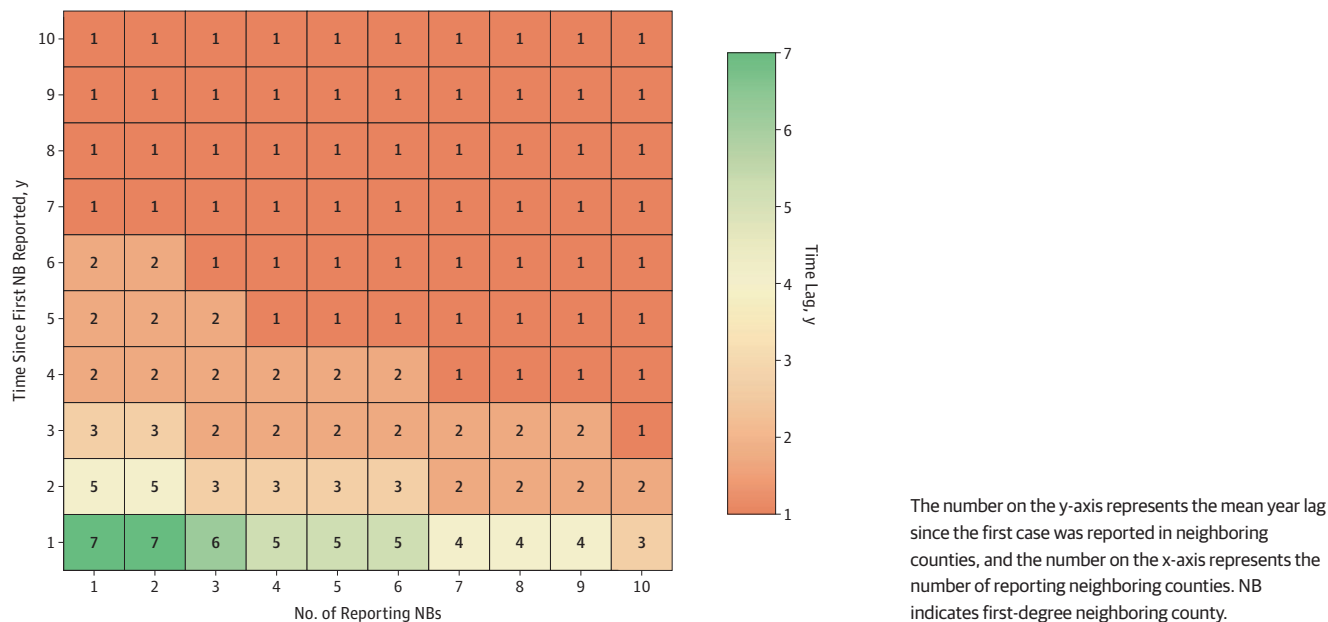


Figure 4. Mean Time Lag of Any County to Report a First Case of Lyme Disease Based on the Reporting Status of Neighboring Counties



but not necessarily where exposure to ticks occurred. Although the factors of neighboring counties were accounted for, travel history beyond first-degree and second-degree neighboring counties could also result in tick exposure and subsequent disease. Such travel could have skewed the estimated associations of the ecological, environmental, and demographic covariates included in the statistical model, particularly given that the outcome variable of the statistical model was reporting at least 1 case of Lyme disease. This limitation was overcome by performing the model selection process 1000 times and randomly excluding 100 counties from the data set in each run (eAppendix in the [Supplement](#)). Therefore, we can assume that Lyme disease cases acquired via travel would have similar associations across the study area if they occurred randomly across the regions. Our findings showed no notable association between noncontiguous counties and the occurrence of at least 1 Lyme disease case. However, we cannot rule out that exposures may have differed during travel in certain regions.

Conclusions

Our predictive model, particularly if it is updated annually and expanded geographically, can enable states and counties to develop more specific Lyme disease prevention and control strategies, including optimizing active surveillance of human disease and ticks to increase early detection. Furthermore, this model will facilitate more targeted communication to and sensitization of the general population and the medical community to the risk of tick exposure and Lyme disease. Our findings can help public health authorities focus surveillance efforts in counties identified as being at high risk of Lyme disease emergence but where cases have yet to be reported.

ARTICLE INFORMATION

Accepted for Publication: December 10, 2019.

Published: March 3, 2020. doi:10.1001/jamanetworkopen.2020.0319

Open Access: This is an open access article distributed under the terms of the [CC-BY License](#). © 2020 Bisanzio D et al. *JAMA Network Open*.

Corresponding Author: Donal Bisanzio, DVM, PhD, Global Health Division, International Development Group, RTI International, 701 13th St NW, Ste 750, Washington, DC 20005 (dbisanzio@rti.org).

Author Affiliations: Global Health Division, International Development Group, RTI International, Washington, District of Columbia (Bisanzio, Reithinger); Division of Epidemiology and Public Health, School of Medicine, University of Nottingham, Nottingham, United Kingdom (Bisanzio, Martello); Earth Institute, Columbia University, New York, New York (Fernández); Department of Ecology, Evolution and Environmental Biology, Columbia University, New York, New York (Fernández, Diuk-Wasser).

Author Contributions: Dr Bisanzio had full access to all of the data in the study and takes responsibility for the integrity of the data and the accuracy of the data analysis. Drs Bisanzio and Fernández contributed equally to the study.

Concept and design: Bisanzio, Reithinger, Diuk-Wasser.

Acquisition, analysis, or interpretation of data: All authors.

Drafting of the manuscript: Bisanzio, Fernández, Reithinger.

Critical revision of the manuscript for important intellectual content: All authors.

Statistical analysis: Bisanzio.

Obtained funding: Diuk-Wasser.

Administrative, technical, or material support: Reithinger.

Supervision: Reithinger.

Conflict of Interest Disclosures: None reported.

Funding/Support: This publication was funded by Cooperative Agreement U01 CK000509-01 from the US Centers for Disease Control and Prevention (CDC) (Drs Fernández and Diuk-Wasser). Dr Fernández was supported by an Interdisciplinary Predoctoral Fellowship from the Earth Institute, Columbia University.

Role of the Funder/Sponsor: The funding sources had no role in the design and conduct of the study; collection, management, analysis, and interpretation of the data; preparation, review, or approval of the manuscript; and decision to submit the manuscript for publication.

Disclaimer: The contents are solely the responsibility of the authors and do not necessarily represent the official views of the CDC or the US Department of Health and Human Services.

REFERENCES

1. Barbour AG, Fish D. The biological and social phenomenon of Lyme disease. *Science*. 1993;260(5114):1610-1616. doi:10.1126/science.8503006
2. Steere AC, Coburn J, Glickstein L. The emergence of Lyme disease. *J Clin Invest*. 2004;113(8):1093-1101. doi:10.1172/JCI21681
3. Killalea ME, Swee A, Lane RS, Briggs CJ, Ostfeld RS. Spatial dynamics of Lyme disease: a review. *Ecohealth*. 2008;5(2):167-195. doi:10.1007/s10393-008-0171-3
4. Nelson CA, Saha S, Kugeler KJ, et al. Incidence of clinician-diagnosed Lyme disease, United States, 2005-2010. *Emerg Infect Dis*. 2015;21(9):1625-1631. doi:10.3201/eid2109.150417
5. Rosenberg R, Lindsey NP, Fischer M, et al. Vital signs: trends in reported vectorborne disease cases: United States and territories, 2004-2016. *MMWR Morb Mortal Wkly Rep*. 2018;67(17):496-501. doi:10.15585/mmwr.mm6717e1
6. Schwartz AM, Hinckley AF, Mead PS, Hook SA, Kugeler KJ. Surveillance for Lyme disease: United States, 2008-2015. *MMWR Surveill Summ*. 2017;66(22):1-12. doi:10.15585/mmwr.ss6622a1
7. Kugeler KJ, Farley GM, Forrester JD, Mead PS. Geographic distribution and expansion of human Lyme disease, United States. *Emerg Infect Dis*. 2015;21(8):1455-1457. doi:10.3201/eid2108.141878
8. Walter KS, Pepin KM, Webb CT, et al. Invasion of two tick-borne diseases across New England: harnessing human surveillance data to capture underlying ecological invasion processes. *Proc Biol Sci*. 2016;283(1832):20160834. doi:10.1098/rspb.2016.0834
9. Pepin KM, Eisen RJ, Mead PS, et al. Geographic variation in the relationship between human Lyme disease incidence and density of infected host-seeking *Ixodes scapularis* nymphs in the eastern United States. *Am J Trop Med Hyg*. 2012;86(6):1062-1071. doi:10.4269/ajtmh.2012.11-0630
10. Hahn MB, Jarnevich CS, Monaghan AJ, Eisen RJ. Modeling the geographic distribution of *Ixodes scapularis* and *Ixodes pacificus* (Acari: Ixodidae) in the contiguous United States. *J Med Entomol*. 2016;53(5):1176-1191. doi:10.1093/jme/tjw076
11. Piedmonte NP, Shaw SB, Prusinski MA, Fierke MK. Landscape features associated with blacklegged tick (Acari: Ixodidae) density and tick-borne pathogen prevalence at multiple spatial scales in central New York State. *J Med Entomol*. 2018;55(6):1496-1508. doi:10.1093/jme/tjy111
12. Connally NP, Ginsberg HS, Mather TN. Assessing peridomestic entomological factors as predictors for Lyme disease. *J Vector Ecol*. 2006;31(2):364-370. doi:10.3376/1081-1710(2006)31[364:APEFAP]2.0.CO;2
13. Centers for Disease Control and Prevention. Lyme disease (*Borrelia burgdorferi*): 2017 case definition. <https://wwwn.cdc.gov/nndss/conditions/lyme-disease/case-definition/2017/>. Accessed December 6, 2019.
14. Cartter ML, Lynfield R, Feldman KA, Hook SA, Hinckley AF. Lyme disease surveillance in the United States: looking for ways to cut the Gordian knot. *Zoonoses Public Health*. 2018;65(2):227-229. doi:10.1111/zph.12448
15. Seltzer EG, Shapiro ED. Misdiagnosis of Lyme disease: when not to order serologic tests. *Pediatr Infect Dis J*. 1996;15(9):762-763. doi:10.1097/00006454-199609000-00003
16. Steere AC, Taylor E, McHugh GL, Logigian EL. The overdiagnosis of Lyme disease. *JAMA*. 1993;269(14):1812-1816. doi:10.1001/jama.1993.03500140064037
17. White J, Noonan-Toly C, Lukacik G, et al. Lyme disease surveillance in New York State: an assessment of case underreporting. *Zoonoses Public Health*. 2018;65(2):238-246. doi:10.1111/zph.12307
18. Waller LA, Goodwin BJ, Wilson ML, Ostfeld RS, Marshall SL, Hayes EB. Spatio-temporal patterns in county-level incidence and reporting of Lyme disease in the northeastern United States, 1990-2000. *Environ Ecol Stat*. 2007;14(1):83-100. doi:10.1007/s10651-006-0002-z
19. Eddens T, Kaplan DJ, Anderson AJM, Nowalk AJ, Campfield BT. Insights from the geographic spread of the Lyme disease epidemic. *Clin Infect Dis*. 2019;68(3):426-434. doi:10.1093/cid/ciy510

20. Chen H, Stratton HH, Caraco TB, White DJ. Spatiotemporal Bayesian analysis of Lyme disease in New York State, 1990-2000. *J Med Entomol*. 2006;43(4):777-784.
21. Lantos PM, Tsao J, Nigrovic LE, et al. Geographic expansion of Lyme disease in Michigan, 2000-2014. *Open Forum Infect Dis*. 2017;4(1):ofw269. doi:10.1093/ofid/ofw269
22. Brinkerhoff RJ, Gilliam WF, Gaines D. Lyme disease, Virginia, USA, 2000-2011. *Emerg Infect Dis*. 2014;20(10):1661-1668. doi:10.3201/eid2010.130782
23. Robinson SJ, Neitzel DF, Moen RA, et al. Disease risk in a dynamic environment: the spread of tick-borne pathogens in Minnesota, USA. *Ecohealth*. 2015;12(1):152-163. doi:10.1007/s10393-014-0979-y
24. Oliver JD, Bennett SW, Beati L, Bartholomay LC. Range expansion and increasing *Borrelia burgdorferi* infection of the tick *Ixodes scapularis* (Acari: Ixodidae) in Iowa, 1990-2013. *J Med Entomol*. 2017;54(6):1727-1734. doi:10.1093/jme/tjx121
25. US Census Bureau. Census regions and divisions of the United States. https://www2.census.gov/geo/pdfs/maps-data/maps/reference/us_regdiv.pdf. Published 2019. Accessed July 5, 2019.
26. von Elm E, Altman DG, Egger M, Pocock SJ, Gøtzsche PC, Vandenbroucke JP; STROBE Initiative. Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) statement: guidelines for reporting observational studies. *BMJ*. 2007;335(7624):806-808. doi:10.1136/bmj.39335.541782.AD
27. Centers for Disease Control and Prevention. Lyme disease: data and surveillance. <https://www.cdc.gov/lyme/datasurveillance/index.html>. Published November 22, 2019. Accessed December 6, 2019.
28. Multi-Resolution Land Characteristics (MRLC) Consortium. <https://www.mrlc.gov/>. Accessed December 6, 2019.
29. University of Wisconsin-Madison SILVIS Lab. Wildland-urban interface (WUI) change 1990-2010. <http://silvis.forest.wisc.edu/data/wui-change/>. Accessed December 6, 2019.
30. NASA Jet Propulsion Laboratory, California Institute of Technology. US releases enhanced shuttle land elevation data. <https://www2.jpl.nasa.gov/srtm/>. Accessed December 6, 2019.
31. Eisen RJ, Eisen L, Beard CB. County-scale distribution of *Ixodes scapularis* and *Ixodes pacificus* (Acari: Ixodidae) in the continental United States. *J Med Entomol*. 2016;53(2):349-386. doi:10.1093/jme/tjv237
32. QGIS [computer program]. Version 3.4. Beaverton, OR: OSGeo; 2019.
33. GRASS [computer program]. Version 7.2. Beaverton, OR: OSGeo; 2019.
34. Burnham KP, Anderson DR. *Model Selection and Multimodel Inference: A Practical Information-Theoretic Approach*. 2nd ed. New York, NY: Springer-Verlag; 2002.
35. Zou KH, Liu A, Bandos A, Ohno-Machado L, Rockette HE. *Statistical Evaluation of Diagnostic Performance: Topics in ROC Analysis*. Boca Raton, FL: Chapman & Hall/CRC; 2016. doi:10.1201/b11031
36. McClure M, Diuk-Wasser M. Reconciling the entomological hazard and disease risk in the Lyme disease system. *Int J Environ Res Public Health*. 2018;15(5):1048. doi:10.3390/ijerph15051048
37. Gasmi S, Ogden NH, Ripoche M, et al. Detection of municipalities at-risk of Lyme disease using passive surveillance of *Ixodes scapularis* as an early signal: a province-specific indicator in Canada. *PLoS One*. 2019;14(2):e0212637. doi:10.1371/journal.pone.0212637

SUPPLEMENT.

eAppendix. Technical Appendix

eFigure 1. Regional Subdivision of the Counties in the Study Area

eFigure 2. Planar Network Used in the Stochastic Model

eFigure 3. Estimated Probabilities of Counties to Report a Lyme Disease Case by 2018 Based on Their Status of Lyme Disease Case Reporting During the 2000-2017 Study Period, by Region.

eReferences.