Citizen science informs human-tick exposure in the Northeastern United States

W. Tanner Porter*, Peter J. Motyka, Julie Wachara, Zachary A. Barrand, Zahraa Hmood, Marya McLaughlin, Kelsey Pemberton and Nathan C. Nieto

Abstract

Background: Tick-borne disease is the result of spillover of pathogens into the human population. Traditionally, literature has focused on characterization of tick-borne disease pathogens and ticks in their sylvatic cycles. A limited amount of research has focused on human-tick exposure in this system, especially in the Northeastern United States. Human-tick interactions are crucial to consider when assessing the risk of tick-borne disease since a tick bite is required for spillover to occur.

Methods: Citizen scientists collected ticks from the Northeastern US through a free nationwide program. Submitted ticks were identified to species, stage, and sex. Blacklegged ticks, *Ixodes scapularis*, were tested for the presence of *Borrelia burgdorferi* sensu lato (s.l.) and hard-tick relapsing fever *Borrelia*. Seasonality of exposure and the citizen science activity during tick exposure was recorded by the citizen scientist. A negative binomial model was fit to predict county level CDC Lyme disease cases in 2016 using citizen science *Ixodes scapularis* submissions, state, and county population as predictor variables.

Results: A total of 3740 submissions, comprising 4261 ticks, were submitted from the Northeastern US and were reported to be parasitizing humans. Of the three species submitted, blacklegged ticks were the most prevalent followed by American dog ticks and lone star ticks. Submissions peaked in May with the majority of exposure occurring during everyday activities. The most common pathogen in blacklegged ticks was *B. burgdorferi* s.l. followed by hard-tick relapsing fever *Borrelia*. Negative binomial model performance was best in New England states followed by Middle Atlantic states.

Conclusions: Citizen science provides a low-cost and effective methodology for describing the seasonality and characteristics of human-tick exposure. In the Northeastern US, everyday activities were identified as a major mechanism for tick exposure, supporting the role of peri-domestic exposure in tick-borne disease. Citizen science provides a method for broad pathogen and tick surveillance, which is highly related to human disease, allowing for inferences to be made about the epidemiology of tick-borne disease.

Keywords: Tick-borne disease, Citizen science, Lyme disease, *Borrelia*, Ticks, *Ixodes*, *Amblyomma*, *Dermacentor*

Background

Tick-borne disease (TBD) diagnoses have steadily increased over the past two decades, along with the spatial distribution of tick-borne pathogens and their associated vectors across the United States [1, 2]. Tick-borne pathogens are transmitted and maintained through a complex cycle of tick vectors, with multiple life stages and a wide diversity of reservoir hosts. For example, *B. burgdorferi* sensu stricto (s.s.), the causative agent of Lyme disease, is transmitted to *Ixodes scapularis* during the larval and nymph life stages, during which the tick is primarily feeding on small mammal hosts. If *B. burgdorferi* was acquired during larval or nymphal feeding, the nymph or adult feeding step allows *B. burgdorferi* to be passed back into naïve rodents, which act as environmental reservoirs for the spirochetes. Once the tick...
progresses to the adult stage, it feeds predominantly on larger mammals, which are considered incidental hosts for *B. burgdorferi* [3]. Humans and domesticated animals are incidental hosts in this cycle, that become infected when tick-borne pathogens spillover from their usual reservoir hosts through the bite of an infected tick [3, 4].

In North America, the majority of TBD are reported from the Northeastern US, a hotspot for Lyme disease, anaplasmosis, babesiosis, and other emerging TBD [5, 6]. TBD are responsible for significant morbidity and health care costs. Lyme disease alone, is estimated to be responsible for approximately 330,000 cases per year and cost an estimated $712 million to $1.3 billion each year in direct medical costs in the US [7, 8]. In the Northeastern US, the black-legged tick, *I. scapularis*, is the most medically important tick species because it is the primary vector of six human pathogens including: *Anaplasma phagocytophilum*, *Babesia microti*, *B. burgdorferi* s.s., *B. miyamotoi*, Powassan virus, and pathogenic *Ehrlichia* spp. [9, 10]. In addition to *I. scapularis*, the American dog tick, *Dermacentor variabilis*, is also common in the Northeast and has been associated with the transmission of pathogenic and non-pathogenic Spotted Fever bacteria (Rickettsiales) and *Francisella tularensis*, the causative agent of tularemia, in the Southern US [9, 11–13]. The lone star tick, *Amblyomma americanum*, has also been collected and reported in several Northeastern states [14], which has been linked to tularemia, pathogenic *Ehrlichia* spp, Heartland virus, *B. lonestari*, and a novel yet severe immune-mediated meat allergy (alpha-gal) [2, 15, 16].

Due to the complexity of TBD cycles, control of disease is a challenging proposition. Currently, human prevention of TBD focuses on decreasing human exposure to ticks and reducing the amount of time a tick is attached to an individual [17]. Previous TBD risk data has utilized studies that incorporate active surveillance, seroprevalence of domesticated animals, landscape dynamics, and citizen science (CS) [18–23]. Active surveillance has been extensively used and utilizes flagging or dragging methods, which measure the density of ticks, allowing inferences to be made about the ecology and evolution of ticks and TBD, and importantly, entomological risk [18, 19]. Active surveillance does not provide information on where and when people were being exposed to ticks, an interaction that is dependent on human activity overlapping with tick biology [24]. CS has been used to answer a variety of questions that range from identifying citizen perceptions of urban environments, plant and animal ecology, classifying images taken from the Hubble Telescope, and utilizing smart phones to investigate health related questions [25–29]. Previously, several CS programs have been utilized to investigate TBD risk across the United States, Southern Canada, the Netherlands, and Finland, however, none have focused on the Northeastern United States [22, 23, 30–34].

Here we present data collected from a CS program that describes the seasonality and characteristics of human-tick interactions in the Northeastern US. These data will inform efforts to decrease the burden of tick-borne disease on the public health system. Additionally, we present county level prevalence of *B. burgdorferi* s.l., and hard-tick relapsing fever *Borrelia* within *I. scapularis* ticks that were found parasitizing humans. Finally, we present correlations between CS-based model predictions of tick exposure and CDC reported human cases of Lyme disease.

**Methods**

**Tick collection**

Ticks were collected by citizen scientists between January 2016 and August 2017 and submitted to a free tick testing campaign (Bay Area Lyme Foundation, https://www.bayarealyme.org/), as described in Nieto et al. [35]. Citizen scientists were recruited through an initial public relations campaign and a public web-site (Bay Area Lyme Foundation, http://www.bayarealyme.org/lyme-disease-prevention/ticktesting/). Additionally, TBD public awareness groups, that were unrelated to the researchers and funders, published advertisements and used social media platforms to promote the program. The citizen scientists participated in the program by completing a submission form and mailing it and a collected tick to the testing facility. No personal-identifying information was recorded (i.e., name, age, gender, etc.).

For this study we analyzed ticks reported to be parasitizing humans and submitted from the Northeastern US, which was defined by the US Census Bureau as Connecticut, Massachusetts, Maine, Vermont, New Hampshire, New Jersey, New York, Rhode Island, and Pennsylvania. Each submission included a questionnaire that detailed the date and location of where the tick was found that included GPS responses, city, county and state responses. For the purpose of this analysis only county and state locations were analyzed. Additionally, the submission included whether or not the tick was questing or biting, host type, and activity the citizen scientist was participating in during the tick exposure. Responses to activity were binned based on the citizen scientist’s free form response into 6 bins: walking or walking pet; outdoor recreation; gardening, yard work, mowing; daily activities; community level recreation; and unknown.

Ticks were identified to species, sex, and stage, based on standard taxonomic keys [36–39]. DNA extractions were performed following manufacturer’s recommendations with a few modifications (DNeasy extraction kit, Qiagen, Valencia, Ca.); ticks were bisected and incubated...
Ixodes scapularis pathogen detection

Ixodes scapularis samples were tested for Borrelia burgdorferi sensu lato (s.l.), and hard-tick relapsing fever Borrelia, using previously published qPCR hybridization primer and probe sets [40]. For the purposes of this paper, B. burgdorferi s.l. refers to the B. burgdorferi s.l. complex, a diverse group of Borrelia spp. that vary in pathogenic potential and includes B. burgdorferi s.s., the causative agent of Lyme borreliosis [41]. Similarly, hard-tick relapsing fever Borrelia refers to spirochetes present in I. scapularis that comprise the Tick-Borne Relapsing Fever (TBRF) complex, such as B. miyamotoi [40]. All reactions were performed in 20 μl reaction volumes with 300 nm primer concentrations and 250 nm probe concentrations (Applied Biosystems, Life Technologies, Carlsbad, CA), and BIORAD SsoAdvanced Universal Supermix (Life Science Research, Bio-Rad, Hercules, CA) was used at the manufacturers recommended concentration. All reactions were performed on a BIORAD CFX96 TOUCH thermocycler (Life Science Research, Bio-Rad, Hercules, CA) with the following steps: 95 °C for 3 min followed by 40 cycles of 95 °C for 10 s and 55 °C for 30 s. Samples were considered positive if they displayed logarithmic expansion prior to 40 cycles and negative controls were included with every analysis.

Analysis

All analyses were performed in the statistical package “R,” version 3.5.0 [42]. All figures were produced with the R-package “ggplot2,” additionally; the package “tidyverse” was used for general data analysis and manipulation [43, 44]. The packages “rgdal,” “maptools,” and “rgeos” were used for the production of county level maps [45–47]. 95% confidence intervals were produced through the proportions test (prop.test) in the base package of “R.” Counties with ten or more submitted I. scapularis ticks were include in the pathogen prevalence maps.

All correlations were computed using a Spearman rank correlation through the correlation, variance and covariance function (cor) in the base package of “R,” Spearman rank correlations were used due to nonlinear relationships. A negative binomial model, using the “MASS” package, was fitted from the CS data set to predict the number of reported human CDC Lyme disease cases in 2016 [48]. County level human CDC cases were retrieved from the CDC through the National Notifiable Disease Surveillance System [49]. County level population data was included from the US Census Bureau decennial census in 2010 [50]. The negative binomial model was chosen due to the nature of the data (count data), distribution of the data (right-skewed distribution), and over dispersion of a Poisson model (p < 0.0001). All Northeastern counties were used in the model, including counties with no I. scapularis submissions. Model fit was assessed through Akaike Information Criterion (AIC) [51], mean absolute error (MAE), Root Mean Square Error (RMSE), and Normalized Root Mean Square Error (NRMSE) was normalized based on the mean of CDC reported Lyme disease cases. Additionally, model fit was assessed based on the location of each county by US Census Division (New England and Middle Atlantic) and by state. These geopolitical lines do not have an impact the ecology of the system, however, variations occur in state reporting guidelines for tick-borne diseases. For example, Massachusetts has changed the reporting guidelines for Lyme disease, resulting in a 20× decrease of reported Lyme disease cases between 2015 and 2016 (2015: 4224 cases, 2016: 198 cases) [5, 49].

Results

Tick submissions

Between January 2016 and August 2017, a total of 3740 submissions were received from the Northeastern US and were reported by the citizen scientist to be biting/questing on a human, with an average of 1.1 ticks per submission (range 1–38), accounting for 4261 ticks. 58 (1%, 58/4261) ticks were unidentifiable at the species level due to decomposition or loss of key morphological characteristics. Submissions were received from 202 of the 217 counties in the Northeastern US (Fig. 1a). Including counties with no submissions, an average of 17 submissions were received from each county (median 9, range 0–124). I. scapularis submissions were received from 196 of the 217 counties with an average of 11 submissions per county (median 5, range 0–85; Fig. 1b).

Tick species

Three tick species were identified from the 4278 ticks that were submitted. I. scapularis accounted for 60% of the total (2574/4261), though this proportion varied across states (e.g., I. scapularis accounted for 37% of submissions in New Jersey (149/404) and 75% in Rhode Island (35/47)). D. variabilis accounted for 33% (1388/4261) of ticks overall, and state submissions ranged from 19% (New York, 189/988) to 51% (New Hampshire, 168/329). A total of 4% (160/4261) of ticks identified were A.
Americanum and 2% (81/4261) were identified as Amblyomma spp. nymphs/larvae. Amblyomma ticks accounted for 0.3% (1/329) in New Hampshire and up to 30% (119/404) in New Jersey (Table 1).

Seasonality of exposure
Ticks were encountered with a bimodal distribution throughout the year; the first and largest pulse began in April and lasted into August with the peak being in May. In 2016, the peak exposure occurred in the last week of
May (23rd week of the year) with 148 submissions per week. In 2017, exposure peaked in the 20th and 22nd week of the year with 306 and 304 submissions per week. The second pulse occurred between October and the end of November. In 2016, this peak occurred in the 43rd and 45th week with 66 and 61 submissions per week (Fig. 2). Similar patterns in seasonal exposure were observed between Middle Atlantic states (New Jersey, New York, 

### Table 1 Percentage of ticks collected by citizen scientists per species and state [% (#/n, 95% CI)]

| State          | I. scapularis | D. variabilis | A. americanum |
|----------------|---------------|---------------|---------------|
| Connecticut    | 56.1 (105/187, 48.7–63.3) | 42.2 (79/187, 35.1–49.7) | 0.5 (1/187, 0–3.4) |
| Massachusetts  | 59.9 (430/718, 56.2–63.5) | 39.1 (281/718, 35.6–42.8) | 0.7 (5/718, 0.3–1.7) |
| Maine          | 63.8 (339/531, 59.6–67.9) | 33.5 (178/531, 29.5–37.7) | 1.3 (7/531, 0.6–2.8) |
| New Hampshire  | 46.2 (152/329, 40.7–51.8) | 51.1 (168/329, 45.5–56.6) | 0.3 (1/329, 0–2) |
| New Jersey     | 36.9 (149/404, 32.2–41.8) | 32.9 (133/404, 28.4–37.8) | 29.5 (119/404, 25.1–34.2) |
| New York       | 69.4 (686/988, 66.4–72.3) | 19.1 (189/988, 16.8–21.8) | 9.3 (92/988, 7.6–11.3) |
| Pennsylvania   | 63.3 (609/962, 60.2–66.3) | 33.8 (325/962, 30.8–36.9) | 1.5 (14/962, 0.8–2.5) |
| Rhode Island   | 74.5 (35/47, 59.4–85.6) | 23.4 (11/47, 12.8–38.4) | 2.1 (1/47, 0.1–12.7) |
| Vermont        | 72.6 (69/95, 62.4–81) | 25.3 (24/95, 17.2–35.4) | 1.1 (1/95, 0.1–6.6) |

**Fig. 2** Seasonality of tick submissions by census division (a), tick species (b), and I. scapularis by life stage (c)
and Pennsylvania) and New England states (Connecticut, Massachusetts, Maine, New Hampshire, Rhode Island, and Vermont) (Fig. 2a).

During the May exposure pulse, both I. scapularis and D. variabilis ticks were submitted at similar frequencies, and A. americanum were submitted in much lower frequencies. During the second peak in 2016 (week 42 and 44), the submissions consisted exclusively of I. scapularis (Fig. 2b).

Overall, ticks were predominantly adults (72%, 3054/4261), followed by nymphs (25%, 1070/4261) and larvae (2%, 75/4261). Adult I. scapularis comprised 60% (1554/2574) of I. scapularis submissions, followed by nymphs (36%, 938/2574) and larvae (3%, 73/2574). 99% (1371/1388) of D. variabilis were adults, followed nymphs (1%, 17/1388). A. americanum adults accounted for 53% (128/241) of A. americanum followed by nymphs (46%, 110/241), and larvae (1%, 2/241).

In 2016, nymph exposure peaked during the 23rd and 24th week, with 60 and 57 submissions per week, accounting for 41% and 45% of weekly ticks. In 2017, nymph exposure peaked in the 22nd week, with 111 submissions per week, accounting for 36% of weekly ticks. During 2017, a nine-week period from week 21 to week 29, nymph exposure was elevated and accounted for 18–57% of total ticks. Larval exposure peaked in the 27th week of 2016 with 16 submissions and during the 27th and 30th week of 2017 with 14 and 15 submissions per week. I. scapularis tick exposure followed these trends, however, the peak of nymphs in June and July was comparable to the exposure of I. scapularis adults in May (Fig. 2c).

Citizen science activity

A large majority of ticks were encountered while the citizen scientists were reportedly engaged in yard work (ex. yard maintenance, gardening, or mowing) (23%, 980/4261), community level outdoor recreation (ex. outdoor field sports, playing outside, picnics, BBQs, etc.) (22%, 36/4261), walking or walking their dog (19%, 800/4261), or during daily activities (ex. sitting inside, driving, cleaning, showering, etc) (16%, 667/4261). In contrast, 11% (486/4261) of ticks were reportedly collected while the citizen scientists were engaged in activities with strong associations to forests (ex. biking, hiking/backpacking, camping, hunting, etc.). Additionally, 9% (392/4261) of samples did not have a clear activity response and were grouped into an unknown category (Fig. 3). Reported CS activity did not vary based on tick species.

Borrelia spp. prevalence in I. scapularis

B. burgdorferi s.l. positive samples were identified from all Northeastern US states in I. scapularis ticks. For counties with ten or more I. scapularis ticks submitted, county prevalence ranged from zero to 57.1% (Tolland, CT) (Fig. 4a). Hard-tick relapsing fever Borrelia positive I. scapularis samples were received from seven Northeastern US states and county prevalence, of counties ten or more I. scapularis submissions, ranged from zero to

![Fig. 3 Citizen scientist reported activity during tick exposure](image-url)
12.5% (Strafford, NH) (Fig. 4b). County level pathogen data for all Northeastern US counties has been included as a supplemental material (see Additional file 1). Additionally, pathogen prevalence was calculated based on the activity the citizen scientist reported. Prevalence ranged from 15 to 26% for B. burgdorferi s.l., 0–3% for hard tick relapsing fever (Table 2).

Citizen science collections and CDC reported cases
County *I. scapularis* submissions were moderately correlated to CDC reported cases in 2016 ($\rho=0.39$). The correlation was significantly improved when counties were compared within states (range 0.27–0.9) (Table 3). The Middle Atlantic states had lower correlations (Average 0.62) than New England states (Average 0.72). In model 1, *I. scapularis* submissions was fitted

---

**Fig. 4** County level prevalence of *B. burgdorferi* s.l. (Lyme group) (a), hard tick relapsing fever (TBRF) (b), counties with < 10 ticks were excluded
to CDC cases, yielding an AIC of 2498. Due to state to state variation in slope, an interaction parameter was fitted between county *I. scapularis* submissions and each state (model 2). This model produced an $\text{AIC} = 2398.2$ ($\Delta \text{AIC} = -99.8$). Finally, county population was added to model 3 producing an $\text{AIC} = 2345.3$ ($\Delta \text{AIC} = -152.7$). Predictions of model 3 produced a MAE = 60.7, RMSE = 101.7, and $\rho = 0.83$, indicating increase in model fit compared to models 1 and 2 (Table 4). Across the entire Northeastern US, predicted cases were highly correlated to CDC reported cases in 2016 (Fig. 5). At the division level, the model performance increased for New England counties with a MAE = 33 versus Middle Atlantic counties which produced a MAE = 73 (Fig. 6a). Model performance separated by state identified reduced model performance in Pennsylvania (MAE = 99) and New York (MAE = 45). Generally, model performance followed a 1:1 linear relationship between predicted and actual CDC reported cases (Fig. 6b).

**Discussion**

Overall, the CS collection program was extremely successful collecting over 4000 ticks, in a short time period (~18 months). Ticks were submitted from 93% of the counties across the Northeastern states. The most commonly submitted tick species was *I. scapularis*, followed by *D. variabilis* and *A. americanum*. Previous county/state level passive surveillance programs have been conducted in several Northeastern states including New York, Maine, and Massachusetts. In our study, 60% of ticks that were submitted from Massachusetts were *I. scapularis*, a significant decrease from a similar pay-for-service tick testing program from Massachusetts, which reported that 95% of ticks were *I. scapularis* [30]. This variation could be explained by a pay-for-service program, which selects for ticks that are traditionally associated with TBD, compared to a free surveillance program which does not have a financial attachment to tick submissions. Additionally, these discrepancies could be explained by changing tick seasonality and/or variation in human exposures from year to year.

*I. scapularis* were responsible for more than 60% of human-tick interactions in Maine, which is similar to a recent passive surveillance program report which documented increasing prevalence of *I. scapularis* between 1989 and 2006 [32, 52]. Overall, *A. americanum* accounted for ~9% of tick bites in New York, which appears to substantiate the increase in human—*A. americanum* interactions [53, 54]. These results corroborate current hypotheses stating that *A. americanum* and *I. scapularis* are undergoing significant range expansion [32, 55, 56].

Regional changes of human-tick exposures can have a significant impact on regional public health concerns. Range expansion of *I. scapularis* (north into Canada and into the Southeastern US), shifts in reservoir and host range, and increased TBD reporting have all attributed to increases in detected TBD such as Lyme disease [57]. Additionally, increasing human interactions with *A. americanum* ticks have the potential to expose residents to pathogenic agents or TBD that are traditionally diagnosed in the Southern US, such as Heartland virus, ehrlichiosis, pathogenic *Rickettsia* spp., *B. lonestari*, and alpha-gal allergies [9, 56]. Overlapping seasonality of human exposure to several tick species adds another layer of complexity to the diagnosis of tick-borne disease.
Table 4  Results of model selection using citizen science to explain CDC reported Lyme disease cases in 2016

| Model | Model parameters | df | Deviance | p   | Null deviance | Residual deviance | AIC | ΔAIC | MAE | RMSE | NRMSE | p   |
|-------|------------------|----|----------|-----|---------------|-------------------|-----|------|-----|------|-------|-----|
| 1     | I. scapularis Submissions | 1  | 262.7   | 9.4 | 253.3         | 2498              | –   | 94.1 | 131.3 | 1.1  | 0.4   |
| 2     | I. scapularis Submissions | 1  | 430.2   | 15.6 | 242.4         | 2398.2            | 99.8| 66.5 | 107.3 | 0.91 | 0.77  |
|       | State            | 8  | 147.9   | 24.4 | 0.002         |                   |     |      |      |      |       |
|       | Submissions: State | 8  | 2345.3  | 30.9 | 2394.0        |                   |     |      |      |      |       |

Fig. 5  Performance of CS based model predictions compared to actual CDC reported Lyme disease cases in 2016. Four elements have been included to improve interpretation: (1) 1:1 black line indicating a perfect model; (2) grey ribbon indicating one standard deviation of the reported Lyme disease cases from the 1:1 line; and (3) points colored based on U.S. Census Division (New England = blue, Middle Atlantic = red)

as expanding vector ranges and overlapping phenology can complicate pathogen detection.

Seasonality of human tick exposure is also critical to consider when attempting to combat tick-borne disease. Previous direct surveillance methods have identified the seasonality of I. scapularis in the northeastern US, with nymphs most active from May to August (peaking in late June) and larvae active from July to September [58–62]. Adult ticks have rarely been included in these analyses; when adults are included, differing seasonality patterns have been observed. Falco et al. [59] identified an increase in adult I. scapularis abundance October through December, while Simmons et al. [62] identified an additional peak in adult activity from April to June. Our CS program indicated that human-tick exposure generally reflected the environmental seasonality of I. scapularis, with adult submissions most closely matching Simmons et al. [62], with
a bimodal distribution. Additionally, these seasonality findings mirror a similar passive surveillance study that was conducted in Massachusetts [30]. Increases in CS submissions likely reflect an overlap in increased tick densities and increased human exposure to ticks, due to increasing outdoor spring/summer activities in the Northeastern US.

The majority of tick exposures occurred just prior to the peak in CDC Lyme disease cases, with over 50% of Lyme disease cases being reported in June and July [8]. Due to their small size, nymphs increase the likelihood of attachment for at least 24–48 h, which is the time it takes for the tick to transmit the pathogen to the host, and Lyme disease cases increase following environmental peaks in the density of nymphs [17, 19, 59]. Analysis
of CDC reported Lyme disease cases between 1992 and 2007 in the Northeast found that on average the Lyme disease season starts between the 20th and 22nd week with a peak in the 27th week and ending between the 34th and 36th week of the year [63]. These findings appear to support the involvement of nymphs as a large public health concern, however, the role of adult ticks in tick-borne disease transmission cannot be eliminated. CS provides a tool to identify when people are coming into contact with ticks, however, this method inherently selects for ticks that people find, thus, the human interactions with nymphs or larvae may be underestimated in our analysis.

A large body of work has focused on characterizing tick densities based on biological communities within residential yards, identifying that wooded areas have the highest densities of ticks followed by ecotones, ornamentals, and lawns [64–66]. Human-tick exposure is dependent on the density of ticks the amount of time an individual spends in each area. CS allowed data to be collected on the activity of the citizen scientist at the believed time of tick exposure. Citizen scientists predominantly reported tick exposure to occur in likely peri-domestic environments or during everyday activities including yard work (23%), community level outdoor recreation (22%) or walking/walking their dog (19%), with the minority of exposure (11%) occurring during activities, such as hiking and hunting. These findings are supported in the literature, which identified Lyme disease infection as the result of peri-domestic tick exposure [19, 67]. These findings emphasize the importance of consistent and daily implementation of personal protective practices against ticks during peak exposure periods.

CS also allows for a low-cost and low-resource method of pathogen surveillance. County level pathogen prevalence can be generated from direct surveillance methods; however, these methods often require a large amount of time and human resources to accurately estimate the prevalence by sampling several study sites. The presented pathogen prevalence are consistent with previously published prevalence data from both CS and active surveillance programs [30, 68]. No difference in Borrelia spp. prevalence was seen between activity types, indicating constant bacterial exposure risk across all tick contacts.

Several other methodologies have been used to model the risk of Lyme disease, including the density of host-seeking nymphs and the density of infected nymphs [69–74]. A pitfall of both of these methodologies is a failure to quantify or include human contact with the vector. CS or crowd sourced surveillance may supplement entomological risk parameters, allowing data from human-tick exposure or public perceptions about exposure to be paired with traditional surveillance techniques [75, 76]. CS provides a proxy to assess the number of human-tick interactions over large and fine geospatial areas. When compared to reported CDC cases in 2016 a simple model using the CS dataset was able to explain a large amount of the variation in CDC cases (Fig. 5). Citizen science could be an invaluable tool to understanding and predicting both nationally notifiable and non-notifiable infectious diseases.

CS has its own pitfalls; this methodology relies on non-structured collection of data that can be skewed by several confounding factors [77]. These factors can include: representativeness to the general population of participating individuals and spatial biases [78]. Our CS program attempted to limit the effects these factors had on the data by utilizing a free program that only financially required the citizen scientists to mail an envelope, with the tick, to the testing facility. Additionally, our limited advertising strategy and availability to the general public through a public website were intended increase accessibility to the program and limit the spatial biases associated with submissions. This can be observed in the parent data set that collected 16,080 ticks from 49 states across the United States [35]. Through the course of the program, outreach groups advertised the program and the program incorporated submissions from citizen scientists who frequently submitted ticks, these factors likely have impacts on the representativeness of the participating population and spatial biases of the data set, however, no information regarding how the individual learned of the program or personal information was collected to assist in identifying and correcting for these submissions. In the future, we intend to continue this free program, perform analysis based on time series models, and tie human tick exposure to ecological determinants. Continuing this program will allow for further investigation of human-tick exposure, potential expansion of ticks and/or TBD, and evolutionary patterns of TBD.

Conclusions

CS tick collection has some shortcomings, these disadvantageous are outweighed by its advantages. CS provided an effective low cost and low resource collection method to identify human-tick exposure. Additionally, these collection methods can be used to identify the prevalence of pathogens in ticks that are parasitizing humans at a geospatial level that would be a large undertaking through traditional methodologies. These data can be used to inform activist groups and the public on current trends and exposure routes of TBD. CS surveillance can be supplemented with traditional collection methods to inform on the evolution and expansion of ticks and TBD. Finally, CS can be used to obtain large data sets that
can be used to assess acarological risk across time and space.

Additional file

Additional file 1. County level prevalence of Lyme and TBRF group bacteria. Supplemental file of County level Lyme and TBRF group bacteria prevalence data from citizen science collected I. scapularis ticks in the Northeastern US.

Abbreviations
TBD: tick-borne disease; CS: citizen science; s.s.: sensu stricto; s.l.: sensu lato; CDC: centers for disease control and prevention; AIC: akaike information criterion; MAE: mean absolute error; RMSE: root mean square error; NRMSE: normalized root mean square error; p: Spearman rank correlation.

Acknowledgements
We would like to thank all of the citizen scientists who made this study possible, all the Nieto Lab personnel for their laboratory support and Dr. Daniel J. Salkeld for his valuable manuscript comments.

Authors’ contributions
WTP, JW, ZAB, ZH, MM, and KP participated in tick identification and lab processing. WTP, PJM, and NCN designed the analysis. WTP performed the data analysis. Results were interpreted by WTP, PJM, and NCN. WTP drafted the manuscript, which was revised by NCN, PJM, JW, ZAB, and ZH. All authors read and approved the final version of the manuscript.

Competing interests
The authors declare that they have no competing interests.

Availability of data and materials
The datasets analyzed in this paper are available in the supporting information or are available from the corresponding author on reasonable request.

Consent for publication
Not applicable.

Ethics approval and consent to participate
Ethics approval for human subjects was not required as no private information or are available from the corresponding author on reasonable request.

Availability of data and materials
The datasets analyzed in this paper are available in the supporting information or are available from the corresponding author on reasonable request.

Consent for publication
Not applicable.

Funding
This research was supported by the Bay Area Lyme Foundation (https://www.bayarealyme.org/) through a grant to NC Nieto. The Bay Area Lyme Foundation had no role in the study design, data collection, decision to publish, or the preparation of the manuscript.

Publisher’s Note
Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Received: 11 January 2019 Accepted: 29 April 2019
Published online: 07 May 2019

References
1. Ostfeld RS, Brunner JL. Climate change and ixodes tick-borne diseases of humans. Philos Trans R Soc Lond B Biol Sci. 2015;370:20140051. http://www.ncbi.nlm.nih.gov/pubmed/25688022.
2. Paddock CD, Lane RS, Staples JE, Labruna MB. Global health impacts of vector-borne diseases. In: Mack A, editor. Global health impacts of vector-borne diseases: workshop summary. Washington, D.C.: National Academies Press, 2016. https://www.ncbi.nlm.nih.gov/books/NBK390439/.
3. Tilly K, Rosa PA, Stewart PE. Biology of infection with Borrelia burgdorferi. Infect Dis Clin North Am. 2008;22:217–34. http://www.ncbi.nlm.nih.gov/pubmed/18452798.
4. Ostfeld RS, Levi T, Keesing F, Ogenfuss K, Canham CD. Tick-borne disease risk in a forest food web. Ecology. 2018;99:1562–73. https://doi.org/10.1002/ect.2386.
5. Adams DA, Thomas KR, Jajosky RA, Foster L, Sharp P, Onweh DH, et al. Summary of notifiable infectious diseases and conditions—United States, 2014. MMWR Morb Mortal Wkly Rep. 2016;63:46. https://www.cdc.gov/mmwr/volumes/63/wr/pdfs/mm6346.pdf. http://www.cdc.gov/mmwr/volumes/63/wr/mm6346.pdf.
6. Adams DA, Thomas KR, Jajosky RA, Foster L, Barozi G, Sharp P, et al. Summary of notifiable infectious diseases and conditions—United States, 2015. MMWR Morb Mortal Wkly Rep. 2017;66:449. https://www.cdc.gov/mmwr/pdfs/mm/6649.pdf.
7. Adron ER, Aucoit J, Lembre KW, Weiner JP. Health care costs, utilization, and patterns of care following Lyme disease. PLoS ONE. 2015;10:1–14.
8. Nelson CA, Saha S, Kugeler KJ, Deloney MJ, Shankar MB, Hindkley AF, et al. Incidence of clinician-diagnosed Lyme disease, United States, 2005–2010. Emerg Infect Dis. 2015;21:1625–31.
9. Eisen RJ, Kugeler KJ, Eisen L, Beard CB, Paddock CD. Tick-borne zoonoses in the United States: Persistent and emerging threats to human health. ILAR J. 2017;58:319–35. https://academic.oup.com ILARjournal/article-lookup/doi/10.1177/105433871774005.
10. Nelder MP, Russell CB, Sheehan NJ, Sander B, Moore S, Li Y, et al. Human pathogens associated with the blacklegged tick Ixodes scapularis: a systematic review. Parasit Vectors. 2016;9:265. http://www.ncbi.nlm.nih.gov/pubmed/27151067.
11. Moncayo AC, Cohen SB, Fritzen CM, Huang E, Yabsley MI, Freye JD, et al. Absence of Rickettsia rickettsii and occurrence of other spotted fever group rickettsiae in ticks from Tennessee. Am J Trop Med Hyg. 2010;83:653–7.
12. Stromdahl EY, Hickling GJ. Beyond Lyme: entomology of tick-borne human diseases with emphasis on the south-eastern United States. Zoonoses Public Health. 2012;59 Suppl 2:48–64. http://www.ncbi.nlm.nih.gov/pubmed/22958250.
13. Stromdahl EY, Jiang J, Vince M, Richards AL. Infrequency of Rickettsia rickettsii in Dermacentor variabilis removed from humans, with comments on the role of other human-biting ticks associated with spotted fever group Rickettsiae in the United States. Vector Borne Zoonotic Dis. 2011;11:969–77. http://www.ncbi.nlm.nih.gov/pubmed/21142953.
14. Springer YP, Jarnevich CS, Monagahan AJ, Eisen RJ, Barnett DT. Modeling the present and future geographic distribution of the lone star tick, Amblyomma americanum (Ixodidae: Ixodidae), in the continental United States. Am J Trop Med Hyg. 2015;93:875–90. http://www.ajtmh.org/content/journals/10.4269/ajtmh.15-0330.
15. Barbour AG, Maupin GO, Teltow GJ, Carter CJ, Pesierman J. Identification of an uncultivable Borrelia species in the hard tick Amblyomma americanum: possible agent of a Lyme disease-like illness. J Infect Dis. 1996;173:403–9. http://www.ncbi.nlm.nih.gov/pubmed/8568302.
16. Platt-Mills TAE, Schuyler AJ, Tripathi A, Commens SP. Anaphylaxis to the carbohydrate side chain alpha-gal. Immunol Allergy Clin North Am. 2015;35:247–60. http://www.ncbi.nlm.nih.gov/pubmed/26928661.
17. Eisen L. Pathogen transmission in relation to duration of attachment by Ixodes scapularis ticks. Ticks Tick Borne Dis. 2018;9:535–42. https://doi.org/10.1016/j.ttbdis.2018.01.002.
18. Eisen L, Eisen RJ. Critical evaluation of the linkage between tick-borne disease risk measures and the occurrence of Lyme disease cases. J Med Entomol. 2018;55:1050–62. https://doi.org/10.1093/medet/may009.
19. Mead P, Hook S, Niesobecki S, Ray J, Meek J, Delorey M, et al. Risk factors for tick exposure in suburban settings in the Northeastern United States. Ticks Tick Borne Dis. 2018;9:319–24. https://doi.org/10.1016/j.ttbdis.2018.01.002.
20. Hendricks B, Mark-Carew M, Conley J. Evaluating the utility of companion animal tick surveillance practices for monitoring spread and occurrence of human Lyme disease in West Virginia, 2014–2016. Geospat Health. 2017;12:582. http://www.ncbi.nlm.nih.gov/pubmed/29239564.
21. McClure M, Duk-Wasser M. Reconciling the Entomological hazard and disease risk in the Lyme disease system. Int J Environ Res Public Health. 2018;15;1. http://www.ncbi.nlm.nih.gov/pubmed/29789489.
65. Maupin GO, Fish D, Zultowsky J, Campos EG, Piesman J. Landscape ecology of Lyme disease in a residential area of Westchester County, New York. Am J Epidemiol. 1991;133:1105–13. http://www.ncbi.nlm.nih.gov/pubmed/2035514.

66. Stafford KC, Magnarelli LA. Spatial and temporal patterns of Ixodes scapularis (Acarina: Ixodidae) in Southeastern Connecticut. J Med Entomol. 1993;30:762–71. http://www.ncbi.nlm.nih.gov/pubmed/8360900.

67. Cronley EK, Carter ML, Mrozinski RD, Ertel SH. Residential setting as a risk factor for Lyme disease in a hyperendemic region. Am J Epidemiol. 1998;147:472–7. http://www.ncbi.nlm.nih.gov/pubmed/9525534.

68. Piedmonte NP, Shaw SB, Prusinski MA, Fierke MK. Landscape features associated with blacklegged tick (Acarina: Ixodidae) Density and tick-borne pathogen prevalence at multiple spatial scales in Central New York State. J Med Entomol. 2018;55:1496–508. https://doi.org/10.1093/jme/tjy111/5055416.

69. Diuk-Wasser MA, Vourc'h G, Cislo P, Hoen AG, Melton F, Hamer SA, et al. Field and climate-based model for predicting the density of host-seeking nymphal Ixodes scapularis, an important vector of tick-borne disease agents in the eastern United States. Glob Ecol Biogeogr. 2010;22:428. https://doi.org/10.1111/j.1466-8238.2010.00526.x.

70. Diuk-Wasser MA, Pepin KM, Eisen RJ, Mead PS, Piesman J, Fish D, 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;86:1062–71. https://doi.org/10.4269/ajtmh.2012.11-0630.

71. Diuk-Wasser MA, Hoen AG, Cislo P, Brinkerhoff R, Hamer SA, Rowland M, et al. Human risk of infection with Borrelia burgdorferi, the Lyme disease agent, in Eastern United States. Am J Trop Med Hyg. 2012;86:320–7.

72. Eisen RJ, Eisen L. Spatial modeling of human risk of exposure to vector-borne pathogens based on epidemiological versus arthropod vector data. J Med Entomol. 2008;45:181–92.

73. Mather TN, Nicholson MC, Donnelly EF, Matyas BT. Entomologic Index for Human Risk of Lyme Disease. Am J Epidemiol. 1995;144:1066–9. http://www.ncbi.nlm.nih.gov/pubmed/9574684.

74. Stafford KC, Carter ML, Magnarelli LA, Ertel SH, Mshar PA. Temporal correlations between tick abundance and prevalence of ticks infected with Borrelia burgdorferi and increasing incidence of Lyme disease. J Clin Microbiol. 1998;36:1240–4. http://www.ncbi.nlm.nih.gov/pubmed/9574684.

75. Bouchard C, Aenishaenslin C, Rees EE, Koff JK, Pelcat Y, Ripoche M, et al. Integrated social-behavioral and ecological risk maps to prioritize local public health responses to Lyme disease. Environ Health Perspect. 2018;126:047008. https://www.ncbi.nlm.nih.gov/pubmed/29671475.

76. Soucy J-PR, Slatculescu AM, Nyiraneza C, Ogden NH, Leighton PA, Kerr JT, et al. High-resolution ecological niche modeling of Ixodes scapularis ticks based on passive surveillance data at the Northern Frontier of Lyme disease emergence in North America. Vector-Borne Zoonotic Dis. 2018;18:235–42. https://doi.org/10.1089/vbz.2017.2234.

77. Langley SA, Messina JP, Moore N. Using meta-quality to assess the utility of volunteered geographic information for science. Int J Health Geogr BioMed Central. 2017;16:1–11.

78. Zhang G, Zhu AX. The representativeness and spatial bias of volunteered geographic information: a review. Ann GIS. 2018;24:151–62.