Journal of Medical Entomology, XX(X), 2020, 1–9
doi: 10.1093/jme/tjaa140
Research

Vector-Borne Diseases, Surveillance, Prevention

Crowdsourced Photographs as an Effective Method for Large-Scale Passive Tick Surveillance

Heather L. Kopsco,[1,2,4] Guang Xu,[3] Chu-Yuan Luo,[3] Stephen M. Rich,[3] and Thomas N. Mather[1,2]

[1]Department of Plant Sciences and Entomology, University of Rhode Island, Kingston, RI 02881; [2]TickEncounter Resource Center, University of Rhode Island, Kingston, RI 02881; [3]Department of Microbiology, University of Massachusetts, Amherst, MA 01003; and [4]Corresponding author, e-mail: hkopsco@uri.edu

Subject Editor: Holly Gaff

Received 19 March 2020; Editorial decision 4 June 2020

Abstract

As tick vector ranges expand and the number of tickborne disease cases rise, physicians, veterinarians, and the public are faced with diagnostic, treatment, and prevention challenges. Traditional methods of active surveillance (e.g., flagging) can be time-consuming, spatially limited, and costly, while passive surveillance can broadly monitor tick distributions and infection rates. However, laboratory testing can require service fees in addition to mailing and processing time, which can put a tick-bite victim outside the window of potential prophylactic options or under unnecessary antibiotic administration. We performed a retrospective analysis of a national photograph-based crowdsourced tick surveillance system to determine the accuracy of identifying ticks by photograph when compared to those same ticks identified by microscopy and molecular methods at a tick testing laboratory. Ticks identified by photograph were correct to species with an overall accuracy of 96.7% (CI: 0.9522, 0.9781; $P < 0.001$), while identification accuracy for *Ixodes scapularis* Say (Ixodida: Ixodidae), *Amblyomma americanum* Linnaeus (Ixodida: Ixodidae), and *Dermacentor variabilis* Say (Ixodida: Ixodidae), three ticks of medical importance, was 98.2% (Cohen's kappa [κ] = 0.9575; 95% CI: 0.9698, 0.9897), 98.8% (κ = 0.9466, 95% CI: 0.9776, 0.9941), and 98.8% (κ = 0.9515, 95% CI: 0.9776, 0.9941), respectively. Fitted generalized linear models revealed that tick species and stage were the most significant predictive factors that contributed to correct photograph-based tick identifications. Neither engorgement, season, nor location of submission affected identification ability. These results provide strong support for the utility of photograph-based tick surveillance as a tool for risk assessment and monitoring among commonly encountered ticks of medical concern.

Key words: medical entomology, public health entomology, surveillance

Tickborne disease is a widespread and growing public health problem in the United States. The Centers for Disease Control and Prevention (CDC) report that between 2004 and 2016, the incidence of tickborne disease doubled, and Lyme disease accounted for 82% of human cases (Rosenberg et al. 2018). Incidence of rickettsial diseases like anaplasmosis and ehrlichiosis has also risen (Dahlgren et al. 2011, Mogg et al. 2020). The documented increase in cases is dependent on numerous known and unknown factors, including improved surveillance and testing, but with one major contributor being the range expansion of disease vectoring ticks. The two main Lyme disease vectoring ticks, blacklegged (deer) ticks (*Ixodes scapularis* Say [Ixodida: Ixodidae]) in eastern and midwestern states and its western blacklegged counterpart (*Ixodes pacificus* Cooley & Kohls [Ixodida: Ixodidae]), increased their recorded county presence by 44.7% in the past 20 yr (Eisen et al. 2016). Numerous sources of surveillance data also suggest that the Lone star tick (*Amblyomma americanum* Linnaeus [Ixodida: Ixodidae]), responsible for transmitting Ehrlichia chaffeensis, various rickettsial species, and triggering a mammalian meat allergy, is spreading north- and west-ward (Springer et al. 2014, Christenson et al. 2017, Sonenshine 2018, Jordan and Egizi 2019, Nelder et al. 2019).

Long-term tick surveillance is an important method for tracking tick population distributions and the relative public health risk to humans and companion animals. Active surveillance, most often conducted using flagging, dragging, or chemical-attractant lure methods, survey ticks directly in the field and can establish important disease risk metrics like density of nymphs and density of infected nymphs/entomological risk index for Lyme disease and other tickborne pathogens (Mather et al. 1996, Pepin et al. 2012, Johnson et al. 2018). However, these methods are often time, labor, and financially limiting.

An alternative sampling strategy is passive surveillance in which tick specimens are submitted directly from the public or partnering agencies like hospitals, wildlife and veterinary clinics, and hunting stations (Lee et al. 2019). While often regarded as a less robust sampling method, passive tick collection provides wide spatio-temporal sampling reach, and facilitates the ability to collect additional attributes regarding the behavior of human, domestic, and wild animal hosts, pathogen infection prevalence, and new disease detection (Ogden et al. 2006; Rand et al. 2007; Tulloch et al. 2017; Nieto et al. 2018; Xu et al. 2018, 2019; Porter et al. 2019). The use of passive submissions has been successfully shown to be a reliable measure of tick abundance, and was also shown to provide an accurate indication of human disease risk (Xu et al. 2016, Nieto et al. 2018, Ripoche et al. 2018).

Much fear exists within the public regarding Lyme disease, which is characterized to be a dangerous, insidious, and difficult-to-diagnose and -treat infection (Aronowitz 1991, Herrington 2004, Auwaerter et al. 2011). This anxiety combined with a confusion regarding which tick species (*I. scapularis* and *I. pacificus* in North America) and stages (nymphs and adults) transmit the disease-causing bacteria have led to the belief that prophylaxis should be taken regardless of the duration and type of tick bite (Mather and Mather 1990, Auwaerter et al. 2011, Kopsco et al., in review). Current Infectious Disease Society of America recommendations suggest that simple species identification and feeding time assessment are enough to establish whether prophylaxis should be administered (Wormser et al. 2006), both of which can be easily provided by photographic surveillance. Additionally, photo-based submission can overcome the limitations to submitting in-hand tick specimens, such as delays related to mailing and processing time, which can put a potential Lyme-disease-carrying tick-bite victim outside the window of potential prophylactic options, or under unnecessary antibiotic administration and worry. However, inconsistencies in photo quality and similar-looking rare and more common ticks can potentially limit the accuracy and reliability of such a method. In the only study reported to date that examined the ability of trained researchers to identify tick species by photograph, Koffi et al. (2017) found that with proper photo quality, trained entomologists could prospectively identify commonly encountered ticks correctly to species with 97.2% accuracy (Kappa [κ] = 0.92, Z = 15.46, P< 0.001).

We sought to retrospectively examine the accuracy of tick identifications in a longstanding photo-based surveillance system by comparing ticks identified by photograph with those identified via microscopy with molecular confirmation. We hypothesized that tick photographs were identified correctly to species by trained tick researchers at least 90% of the time. We expected that tick life stage, the region of submission, and researcher-expressed uncertainty in an identification would most strongly contribute to a correct tick identification. Specifically, we anticipated that nymph stage ticks would be incorrectly identified more often than adult stage ticks, submissions from regions that had a greater diversity of tick species that appear visually similar (e.g., *Dermacentor* spp.) would be incorrectly identified more often than those from locations with visually distinct species, and that submissions with poor images (e.g., Fig. 1B) for which researchers could not confidently determine a species would be incorrectly identified more often than those submissions with quality images.

Methods

Data Collection

We compared photograph submissions of tick species sent to TickSpotters, a passive surveillance program at the University of Rhode Island's TickEncounter Resource Center (www.tickencounter.org/tickspotters), to matched in-hand tick specimens sent to the TickReport program at the University of Massachusetts' Laboratory of Medical Zoology (www.tickreport.com) from 2015 to 2017. Included within the TickSpotters program response email is a suggestion for tick testing, and a link to the TickReport program submission page. Due to this connection, it was assumed that there would be a certain amount of overlap in tick submissions to both programs.

The TickReport program provides tick identification and pathogen testing services. They collect mailed tick specimens and record a preliminary species and stage identification during microscopic examination of both the dorsal and ventral surfaces of the tick. Identification to species level is confirmed using real-time quantitative polymerase chain reaction, and equivocal results are resolved with DNA sequencing.

For tick surveillance purposes and assistance in photograph identification, the TickSpotters program asks participants to include with their photograph the date when the tick was discovered, the most reasonable zip code of encounter, whether there was any travel history in the previous 5 d, and whether the tick was found on a person, pet, or loose in the home. A section for more elaboration on the encounter is also included on the submission form. Instructions are provided in the submission form to use top lighting and a light-colored background to ensure a well-lit and focused photograph of the tick that allows the evaluator to visualize key anatomical landmarks like number of legs, dorsal scutum, and mouthparts (Fig. 1). Submitters are encouraged to also include a common size reference like a coin, ruler, or other recognizable item.

Submissions are reviewed by a small team of formally trained career medical entomologists, or advanced students (referred to in the program as 'TickExperts') who are trained in tick identification, anatomy, and biology and directly supervised by these entomologists. Upon identifying the tick to species and stage, TickSpotters TickExperts also estimate the likely duration of feeding by comparing the submitted tick to a visual tick growth tool available on the TERC website (https://tickencounter.org/tick_identification/tick_growth_comparison); feeding duration estimates are based on the scutal index (Yeh et al. 1995, Falco et al. 2018) and used to provide a likelihood of transmission risk assessment in the responding e-mail. If a tick cannot be identified to species via photograph due to poor photo quality (i.e., too blurry, dark, or poorly positioned to identify the key anatomical landmarks previously mentioned), an email is sent requesting an additional image, and the tick is classified as 'Unknown' unless a more definitive identification is made upon receipt of a higher quality photograph. TickSpotters data collection is approved and overseen by the University of Rhode Island Institutional Review Board.

Comparing Datasets

We examined the distribution of overlapping submissions between TickSpotters and TickReport to identify the most common region of submission, the species and stage of ticks submitted, the engorgement status, and the season of submission. A matched TickSpotters photo/TickReport in-hand submission was defined as any entry within the time period that had the same email address and was submitted to both programs within 48 h to conservatively match tick photos to mailed-in tick specimens. However, we recognize that these criteria still allowed for the possibility that the photograph sent to TickSpotters was not of the same tick that was sent to TickReport, resulting in a potential mismatch of photo and in-hand specimen results. Cross-referencing of matched submissions was conducted as carefully as possible, but there may have been an insignificant number of false negatives or false positives.

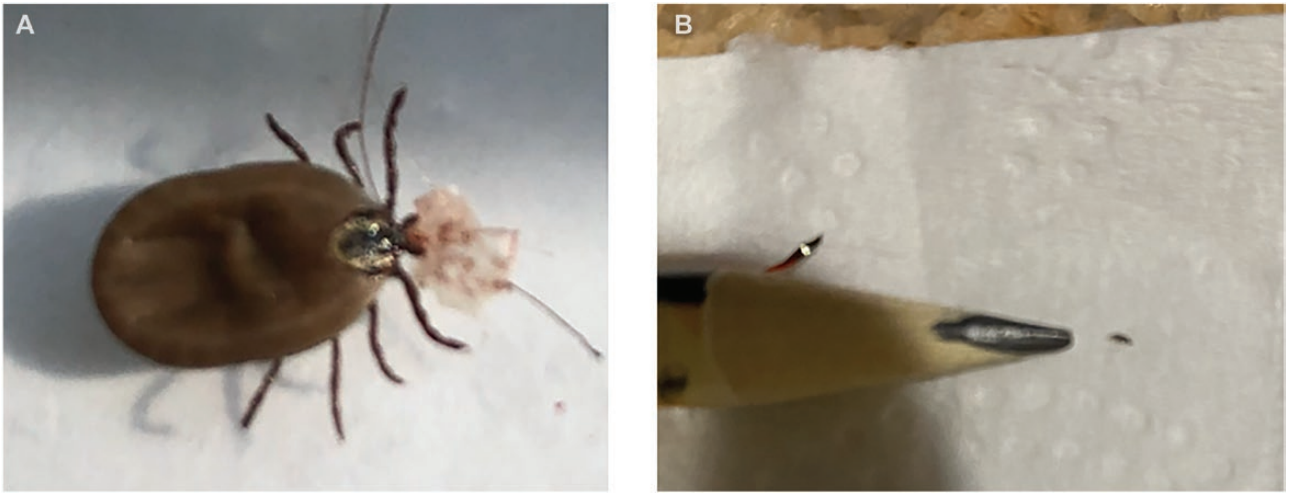

Fig. 1. Examples of photographs submitted to the TickSpotters program. (A) A photograph submission that follows the guidance provided in our submission instructions. (B) A photograph that received an 'Unknown' designation and request for clearer photograph.

For TickSpotters accuracy assessment, photo-based identification was judged incorrect when the TickSpotters-determined genus, species, or both did not match the TickReport record. When a photo was too blurry or did not include the proper lighting to be identified, an additional photograph was requested. If a clearer photograph was not sent or the specimen was still unidentifiable by TickSpotters researchers, these entries were called 'Unknown' and classified as 'Incorrect' identifications due to a 'failure' for the system to be able to identify from a photograph. If a TickSpotters researcher made a 'most likely' identification call but was not entirely certain without further photographic clarification (denoted in the TickSpotters response with question marks), the identification included a binary 'Certain' or 'Uncertain' score, and this variable was included as a separate covariate in the model.

Statistical Analysis

Using TickReport identification as the reference value and the TickSpotters identification as the test value, we performed a classification accuracy analysis to test the hypothesis that over 90% of the TickSpotters identifications (test) matched the TickReport identifications (reference) by establishing a confusion matrix to calculate the overall proportion of correct photograph identification to incorrect, as well as sensitivity, specificity, and positive and negative predictive values. We then performed the same test for *I. scapularis*, *A. americanum*, and *Dermacentor variabilis* Say submissions individually to establish the level of accuracy for these three species of particular medical concern and range expansion out of the entire dataset. We used Cohen's kappa statistic to evaluate the inter-rater reliability between TickReport and TickSpotters tick identifications. Using a scale of 0–1, Cohen's kappa and the corresponding CI report how well the test classifier (i.e., TickSpotters identification) is performing against the reference value (i.e., TickReport identification) beyond random chance (Landis et al. 1977). In other words, Cohen's kappa denotes the level of agreement between two categorical groups, with a value of zero indicating no agreement, and values close to 1 suggest perfect agreement. We opted for this method over simply looking at the percentage correct for each class of tick species because classification accuracy can be misleading in the event of unequal numbers of observations, as was the case for these data. A confusion matrix using Cohen's kappa incorporates and accounts for errors such as false negative and false positives in the performance prediction (Landis et al. 1977).

To model the probability of a correct tick identification by photograph, we fit a logistic regression, a generalized linear model (GLM) that connects a binary outcome to a set of predictors using a logit link function (equation 1). Out of a set of eight covariates we investigated, the possible predictive factors involved with a correct tick photograph identification. These covariates were: encounter season, tick species, tick life stage, engorgement or feeding status (days), uncertainty (binary measure) of researcher in the identification based on the photograph quality, state of encounter (fine spatial scale), region of encounter (broad spatial scale), and host (either loose and wandering, on a pet, or on a person).

$$log\left(\frac{P\,(X)}{1-P\,(X)}\right) = \quad \beta_0 + \quad \beta_1 \quad X_1 \ldots \quad \beta_8 X_8 \qquad (1)$$

where the log odds, or probability (***P***), that a tick is identified correctly (***X***), is equal to the coefficients ($\beta_0 + \ \beta_1 ... \beta_8$) multiplied by the values of the eight covariates as mentioned above. ($X_1 \ldots X_8$). The log odds increase or decrease according to a one-unit change in ***X***.

Training and testing datasets were established by randomly selecting 400 observations, and models were evaluated considering several combinations of parameters using Akaike's Information Criterion (AIC) and Bayesian Information Criterion (BIC), commonly used model selection criterion. Upon selection of the top optimal models using information criterion, we performed additional model fitting using McFadden's pseudo R, and Hosmer and Lemeshow goodness-of-fit (GOF) test for risk prediction and calculated a receiver operator curve (ROC) on the testing dataset. Computation and analysis were performed in R version 3.6.1/ RStudio version 1.2.1335.

Results

Between January 2015 and December 2017, 816 overlapping records of tick photos and in-hand tick specimens were submitted to TickSpotters (3.83% of all submissions during this period) and TickReport (3.50% of all submissions during this period), respectively. Of the total submissions during this time period, TickReport received 74.7% *Ixodes* spp., 14.5% *Dermacentor* spp., 9.7% *Amblyomma* spp., and 1.1% other tick species or specimens. TickSpotters pool of submissions comprised 34.8% *Ixodes* spp., 34.8% *Dermacentor* spp., 17.0% *Amblyomma* spp., 2.3%

Rhipicephalus spp., while other specimens that included unidentifiable ticks, rare ticks, and non-tick arthropods made up 11.1% of all photo submissions (Fig. 2). Submissions compared were sourced predominantly from northeastern and mid-Atlantic states (67% were from Massachusetts, New York, Pennsylvania, Rhode Island, New Jersey, Maryland, and Connecticut collectively).

Tick species represented commonly encountered North American ticks of human and domestic animal concern, with *I. scapularis* representing 68.6% of ticks submitted to both services, followed by *D. variabilis* (13.9%), *A. americanum* (13.4%), *I. pacificus* (2.20%), *Dermacentor andersoni* Stiles (0.61%), *Dermacentor* spp. (0.49%), *Ixodes* spp. (0.36%), *Amblyomma* spp. (0.24%), *Rhipicephalus sanguineus* Latrielle (0.12%), and less common tick species (0.12%) (Table 1). While TickSpotters received a broader array of species submission, these proportions matched those received by TickReport during the same time period. Adult ticks comprised 74.2% of submitted specimens, while nymphs (23.0%) and larvae (0.85%) were considerably less commonly submitted. Stage could not be determined for 1.96% of ticks submitted by photograph (Table 1). Ticks submitted were mostly unfed (median engorgement = 1.5 days), but demonstrated a bi-modal distribution with peaks at <1 d (unfed), and 3 d (partially fed), and a maximum fed time at 8 d. *Ixodes scapularis* comprised more than half of the reported ticks found to be feeding for 3, 3.5, and 4 d. Ticks submitted to both services were predominantly found on humans (90.6%), followed by pets (6.7%), and those found loose and wandering (2.7%) (Table 1). Ticks were submitted mostly in the spring (36.2%), with summer (28.7%) and fall (27.9%) having nearly identical submission rates, and winter (7.2%) being the least common season for submissions. *Ixodes scapularis* was the most common overlapping tick submitted during all seasons, and this species made up nearly all of tick submissions in the fall and winter months.

TickSpotters Photo Identification Accuracy Analysis

Overall, the TickSpotters researchers correctly identified ticks to species from photos with 96.7% accuracy (789/816; CI: 0.9495, 0.9761; $P < 0.001$). The 27 incorrect identifications included six submissions with poor photo quality that were determined as 'Unknown' (Table 2). When examining species of particular disease concern for humans and pets, identification accuracy was highest for *A. americanum* (98.6%; CI: 0.9761, 0.9933; $P < 0.001$) (Fig. 3), followed by *Dermacentor* species (Fig. 5) including *D. variabilis*, *D. occidentalis* Marx, and *D. andersoni* (98.6%; CI: 0.9761, 0.9933; $P < 0.001$) and *I. scapularis* (97.9%; CI: 0.9669, 0.9878; $P < 0.001$) (Fig. 4; Table 3).

The most commonly misidentified tick species occurred in order of their proportional representation, with *I. scapularis* misidentified by TickSpotters the most (10 misidentified ticks; 1.78% of all *I. scapularis* received), and most frequently identified as 'Unknown' (Table 2). Both *A. americanum* (six misidentified, 5.50% of all *A. americanum* received) and *D. variabilis* (five misidentified, 4.42% of all *D. variabilis* received) were most often misidentified as *I. scapularis* (Table 2).

Logistic Regression Model for Covariates Associated With Identifying a Tick by Photograph

Based on the distribution of the data and binary outcome (i.e., correct or incorrect tick identification), a logistic regression was fitted. After examining models containing several combinations of parameters and comparing them to the null (intercept only) model, the

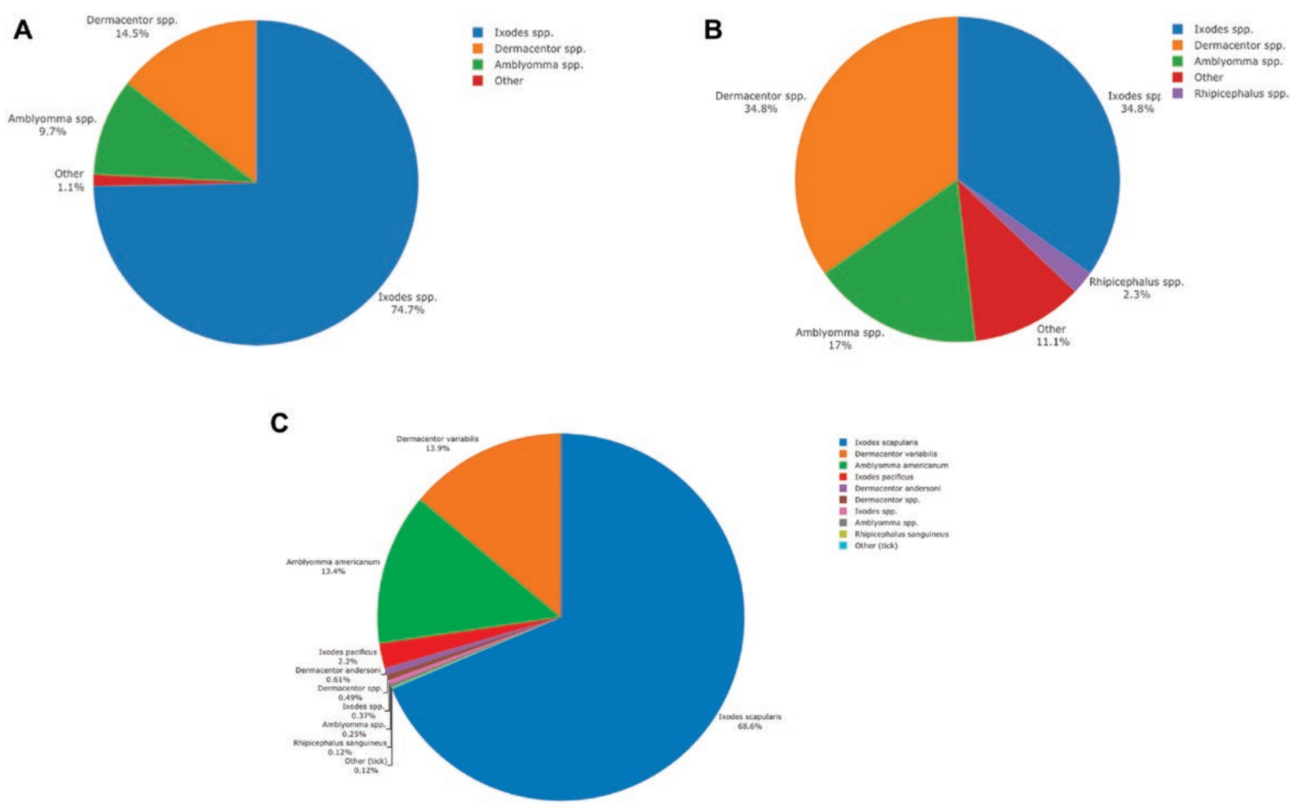

Fig. 2. Proportional tick species contribution to the overall submission pool for both TickReport and TickSpotters during the period from which the overlapping sample was taken (1 January 2015 through 31 December 2017). (A) TickReport (n = 23,379), (B) TickSpotters (n = 21,287), (C) overlapping matched sample (n = 816) of ticks that was sent to both TickReport and TickSpotters, and was used to test the accuracy of TickSpotters identifications.

Table 1. Demographics of overlapping TickSpotters and TickReport submissions from January 2015 to December 2017 (n = 816)

Variable	Categories	Total overlapping sample (%)
Region of residence	Northeast (ME, NH, VT, MA, RI, CT, NY, PA, NJ)	576 (70.4%)
	Southeast (DC, DE, MD, VA, WV, NC, SC, GA, FL, AL, TN, KY, MS, LA, AR)	120 (14.8%)
	Midwest (MO, OH, IN, MI, IL, WI, MN, IA, KS, NE, SD, ND)	83 (10.3%)
	Pacific (WA, OR, CA)	28 (3.4%)
	Mountain (MT, CO, WY, UT, ID, NV)	5 (0.6%)
	Southwest (OK, TX, NM, AZ)	4 (0.5%)
Tick species submitted	*Ixodes scapularis*	560 (68.6%)
	Dermacentor variabilis	113 (13.9%)
	Amblyomma americanum	109 (13.4%)
	Ixodes pacificus	18 (2.20%)
	Dermacentor andersoni	5 (0.61%)
	Dermacentor spp.	4 (0.49%)
	Ixodes spp.	3 (0.36%)
	Amblyomma spp.	2 (0.24%)
	Rhipicephalus sanguineus	1 (0.12%)
	Other (tick)	1 (0.12%)
Stage of tick submitted	Adult	606 (74.2%)
	Nymph	187 (23.0%)
	Unknown	16 (1.96%)
	Larva	7 (0.85%)
Tick engorgement (days)	Unfed to less than 1 d fed	257 (29.3%)
	1–3 d fed	372 (46.9%)
	3.5–5 d fed	173 (21.6%)
	>5 d	14 (2.2%)
Tick host	Human	740 (90.6%)
	Domesticated animal	55 (6.7%)
	Loose and wandering	21 (2.7%)
Season of submission	Spring (Mar. 1–May 31)	296 (36.2%)
	Summer (June 1–Aug. 31)	235 (28.7%)
	Fall (Sept. 1–Nov. 30)	228 (27.9%)
	Winter (Dec. 1–Feb. 28)	59 (7.2%)

Table 2. Ticks incorrectly identified by photograph through the TickSpotters program (n = 28) from January 2015 to December 2017

Tick species	Number of missed identifications (% total of species)	Percent of overall sample (n = 816)	Most commonly misidentified as
Ixodes scapularis (n = 560)	10 (1.78%)	1.22%	Unknown
Amblyomma americanum (n = 109)	6 (5.50%)	0.73%	*Ixodes scapularis*
Dermacentor variabilis (n = 113)	5 (4.42%)	0.61%	*Ixodes scapularis*
Amblyomma spp. (n = 2)	2 (100%)	0.25%	*Amblyomma americanum*
Rhipicephalus sanguineus (n = 1)	1 (100%)	0.12%	*Amblyomma americanum*
Ixodes pacificus (n =18)	1 (5.55%)	0.12%	*Rhipicephalus sanguineus*
Ixodes spp. (n = 3)	1 (33.3%)	0.12%	*Rhipicephalus sanguineus*
Dermacentor spp. (n = 4)	1 (25%)	0.12%	*Dermacentor variabilis*
Other tick (n = 1)	1 (100%)	0.12%	*Rhipicephalus sanguineus*
Tick life stage	**Number of missed identifications (% of total stages)**	**Percent of overall sample (n = 816)**	**Most commonly misidentified as**
Adult (n = 606)	3	0.37%	Nymph
Nymph (n = 187)	0	0	Not applicable
Larva (n =7)	0	0	Not applicable

model incorporating tick species and tick stage as predictors was selected as optimal using both AIC and BIC (Table 4).

The optimal model (tick species and life stage) produced both the lowest AIC and BIC scores relative to the other models. Adding interactions between the two predictors within this model did not improve model fit. The Wald test revealed that both tick species and stage are significant covariates within the top model, and while McFadden's pseudo R for this model was the smallest among models assessed (0.32; range = 0.32–0.46), an R^2 value between 0.2 and 0.4 indicates excellent fit (Domencich and McFadden 1975). Within 20 iterations, there were no statistically significant values ($P < 0.05$) returned in Hosmer and Lemeshow GOF tests, indicating that there was no significant difference between the observed data and the values that were predicted by the model. The ROC for the species + stage model reflected an area under the curve (AUC) of 0.919, which was the largest AUC of any of the models examined

Fig. 3. Specificity and sensitivity analyses of TickSpotters photo-based identification of *A. americanum* encounters (*n* = 109). The number of false positives, true positives, false negatives, and true negatives in relation to the overall abundance of *A. americanum* in the overlapping submission sample (absent *A. americanum* = another tick species vs present = *A. americanum*) are shown as identified by TickReport. Each tick included in the analysis is represented by one of the symbols in the legend. Axes represent the percent of the total sample of all ticks. Proportions of absent (Tn) vs present (Tp) are along the x-axis, while proportions of Fn vs Tp, and Fp vs Tn can be found on the y-axis. (Produced in shinyApp by Allen et al. 2017.)

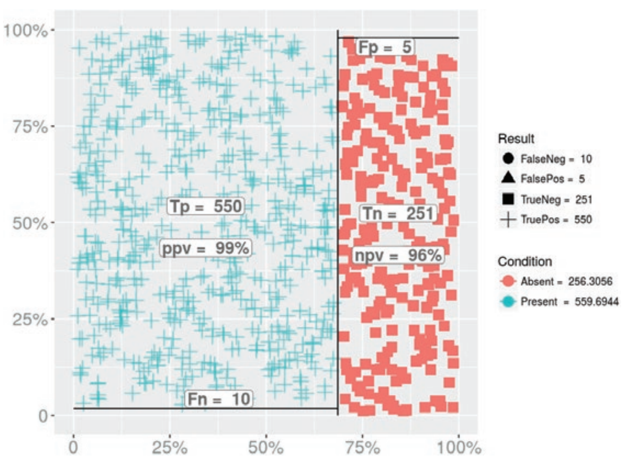

Fig. 4. Specificity and sensitivity analyses of TickSpotters photo-based identification of *I. scapularis* encounters (*n* = 560). The number of false positives, true positives, false negatives, and true negatives in relation to the overall abundance of *I. scapularis* in the overlapping submission sample (absent *I. scapularis* = another tick species, vs present = *I. scapularis*) are shown as identified by TickReport. Each tick included in the analysis is represented by one of the symbols in the legend. Axes represent the percent of the total sample of all ticks. Proportions of absent (Tn) vs present (Tp) are along the x-axis, while proportions of Fn vs Tp, and Fp vs Tn can be found on the y-axis. (Produced in shinyApp by Allen et al. 2017.)

(Fig. 6). These results suggest that the two parameters included in the model demonstrate strong predictive power for explaining the factors influencing tick identifications by photograph.

Discussion

Community science, or when scientists engage members of the public in data collection and other aspects of research projects, is

Table 3. TickSpotters photograph tick identification confusion matrix accuracy analysis for *A. americanum* (*n* = 109), *I. scapularis* (*n* = 560), and *Dermacentor* spp. (*n* = 113)

	Amblyomma americanum (*n* = 109)	*Ixodes scapularis* (*n* = 560)	*Dermacentor* spp. (*n* = 113)
Accuracy	0.9866	0.9792	0.9866
95% CI	(0.9761, 0.9933)	(0.9669, 0.9878)	(0.9761, 0.9933)
P value	<0.001	<0.001	<0.001
Kappa	0.9416	0.9522	0.9472
Sensitivity	0.9364	0.9786	0.9512
Specificity	0.9944	0.9806	0.9928
Positive predictive value	0.9626	0.9910	0.9590
Negative predictive value	0.9902	0.9547	0.9914
Prevalence	0.1345	0.6846	0.1491
Detection rate	0.1259	0.6699	0.1430
Balanced accuracy	0.9654	0.9796	0.9720
AUC	0.98	0.98	0.98

an increasingly popular method for vector-borne disease surveillance. Across Europe, the Mosquito Alert program, designed to employ the public to help track the invasion of Asian tiger mosquitoes (*Aedes albopictus*), found that engaging citizens for data collection can provide an economical and scalable method for collecting surveillance data (Palmer et al. 2017). Similarly, utilizing digital and app-based programs have allowed scientists to recruit citizens in monitoring tick activity in Lyme disease-emergent areas like parts of southern Canada (Hines and Sibbald 2015, Lewis et al. 2018). In the United States, Nieto et al. (2018) established a nationwide tick and pathogen surveillance that identified new foci of disease and tick distribution, and inexpensively monitored tick risk and encounter demographics throughout highly tick and Lyme-endemic northeastern states (Porter et al. 2019). The TickReport program is one of the few services that provides tick species and pathogen testing results to the public in addition to aggregating and interpreting the passively collected data (Xu et al. 2016). Each of these methods, however, relies on the collection of and processing of specimens which can leave participants who encountered a biting tick often waiting several days for a species confirmation.

Similar to a prospective surveillance analysis (Koffi et al. 2017), we found that tick photographs that include adequate lighting and focus to visualize key anatomical features, including the scutum, palps, number of legs, and degree of engorgement (i.e., scutum:body ratio), can be correctly identified to species by those properly trained in tick anatomy and identification more than 90% of the time. According to the optimal model, TickSpotters identification accuracy is dependent on ticks being among the nine most commonly encountered species and life stage, and was not confounded by region of the country, season of submission, or engorgement of the photographed tick. While these findings support the reliability of tick identification by photograph, modeling results underscore the importance of including size references in tick photographs, as nymphs and larvae were significantly less likely to be identified correctly than adult stage ticks. Knowledge of travel history also is critical for a photographic surveillance system to remain reliable given the similar-looking tick species that occur in different parts of the country (e.g., *I. scapularis* and *I. pacificus*) and is why TickSpotters includes a question regarding travel on its submission site. For example, in 2016 after the United States eased travel restrictions to Cuba, TickSpotters started getting out-of-season submissions of partially engorged nymph stage

Table 4. GLM summaries for the five top performing models predicting tick identification accuracy by photograph

	Spp + Stg	Spp + Stg + Uncert	Spp + Stg + Reg	Spp + Stg + Reg + Sea	Spp +Stg + Uncert + Reg
(Intercept)	**7.61 *****	7.90 ***	8.09 ***	7.77 ***	8.38 ***
	(1.52)	(1.68)	(1.69)	(1.80)	(1.84)
Species					
Dermacentor variabilis	**−4.25 ****	−4.56 **	−4.73 **	−5.12 **	−5.05 **
	(1.45)	(1.63)	(1.68)	(1.87)	(1.84)
Amblyomma americanum	**−0.00**	−0.10	0.02	−0.35	−0.08
	(1.29)	(1.31)	(1.47)	(1.72)	(1.51)
Ixodes pacificus	**13.11**	13.16	−7.05	−6.87	−6.83
	(2846.21)	(2801.27)	(4511.27)	(33098.78)	(4485.81)
Other	**−6.32 *****	−6.65 ***	−25.43	−30.98	−25.72
	(1.77)	(1.90)	(3538.43)	(26287.60)	(3534.97)
Amblyomma spp.	**−22.43**	−22.33	−58.40	−72.90	−58.29
	(6522.64)	(6522.64)	(9879.78)	(73053.23)	(9878.54)
Stage					
Nymph	**−3.75 ****	−4.15 **	−4.17 **	−5.07 **	−4.56 **
	(1.35)	(1.55)	(1.46)	(1.75)	(1.66)
Larva	**−6.51 *****	−7.15 **	−7.12 ***	−25.00	−7.71 **
	(1.86)	(2.22)	(2.09)	(3350.62)	(2.40)
Uncertainty					
Uncertain		0.80			0.73
		(1.36)			(1.37)
Region					
Southeast			−0.43	0.17	−0.37
			(1.33)	(1.48)	(1.35)
Midwest			0.54	0.17	0.65
			(1.52)	(1.54)	(1.60)
Pacific			20.05	25.00	19.89
			(3538.43)	(26287.60)	(3534.97)
Mountain			35.91	45.78	35.91
			(7420.60)	(54899.05)	(7418.95)
Southwest			14.10	19.08	14.30
			(3531.03)	(25282.85)	(3520.36)
Season					
Summer				1.86	
				(1.23)	
Autumn				35.47	
				(4562.15)	
Winter				14.63	
				(8468.02)	
AIC	**65.15**	66.77	70.47	68.39	72.15
BIC	**96.84**	102.42	121.96	131.76	127.61
Log likelihood	**−24.57**	−24.38	−22.23	−18.19	−22.08
Deviance	**49.15**	48.77	44.47	36.39	44.15
No. obs.	**388**	388	388	388	388

The model incorporating the covariates of tick species and tick stage was the most predictive of tick identification by photograph and chosen as the optimal model. Other = unidentifiable ticks, rare ticks, and non-tick arthropods. Coefficients are listed with the SEs in parentheses. Optimal model results are bolded. Spp = species; Stg = stage; Sea = season; Reg = region; Uncert = uncertainty.

****P* < 0.001, ***P* < 0.01.

Amblyomma ticks from states where *Amblyomma* are not known to occur. In four cases, questioning travel history revealed recent returns from Cuba, where *Amblyomma cajennense* is commonly encountered. Querying recent travel history can also be critical for tracking tick trends; one Colorado TickSpotter found a 3- to 4-d engorged adult female Asian longhorned tick (*Haemaphysalis longicornis*) on his dog, but we learned he had been in New Jersey with the dog 3 d before he found the tick.

Submission of digital photographs of ticks allows rapid (24 h or less) responses containing public health assurance of likely disease risk, and best next steps for appropriately managing the current tick encounter and helping prevent future tick bites. Among the most commonly encountered ticks in North America, disease risk and the pathogens transmitted by tick bite vary greatly. While blacklegged and western blacklegged ticks are competent vectors of Lyme disease bacteria, other tick species, including American dog ticks and Lone Star ticks, are not (Mather and Mather 1990, Stromdahl et al. 2018). Unfortunately, this fact is not widely appreciated or acknowledged by the general public, who frequently are anxious about risk of a Lyme disease infection following any tick encounter (Auwaerter et al. 2011, Fogel and Chawla 2017). If not quickly informed about the correct tick identification and possible disease associations, citizens will seek, and medical professionals will potentially treat tick bites unnecessarily with antibiotics. Given the increasing problem of antibiotic resistance (Ventola 2015), avoiding unnecessary prophylactic treatment of any length is critically important. Further, properly identifying

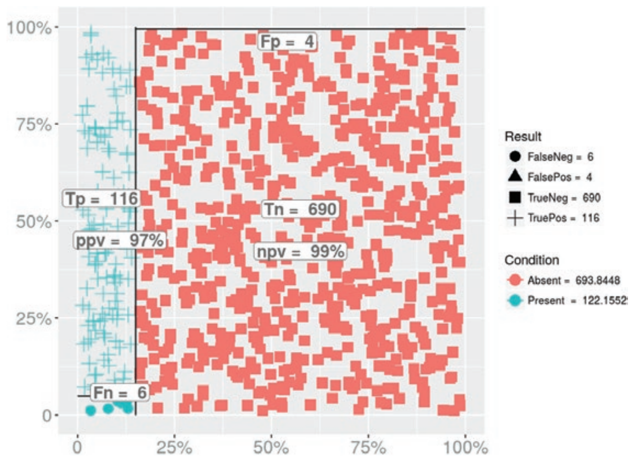

Fig. 5. Specificity and sensitivity analyses of TickSpotters photo-based identification of *Dermacentor* spp. encounters (n = 113). The number of false positives, true positives, false negatives, and true negatives in relation to the overall abundance of *I. scapularis* in the overlapping submission sample (absent *Dermacentor* spp. = another tick species, vs present = *Dermacentor*) are shown as identified by TickReport. Each tick included in the analysis is represented by one of the symbols in the legend. Axes represent the percent of the total sample of all ticks. Proportions of absent (Tn) vs present (Tp) are along the x-axis, while proportions of Fn vs Tp, and Fp vs Tn can be found on the y-axis. (Produced in shinyApp by Allen et al. 2017.)

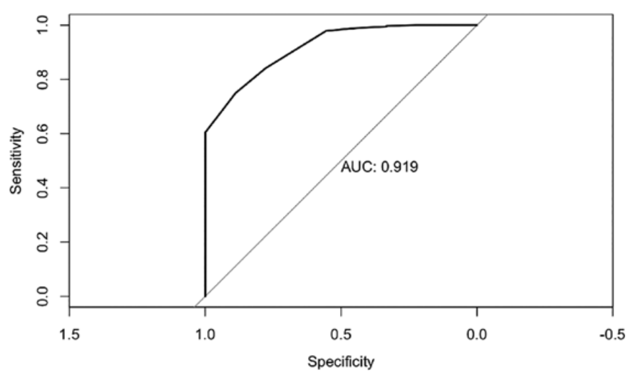

Fig. 6. Receiver operator curve (sensitivity (true positives) vs 1-specificity (false positives)) for the optimal model incorporating species, stage, season, and region as predictors for a correct tick identification by photograph, reporting an AUC of 0.919, suggesting that this model predicts the ability to identify a tick with excellent sensitivity.

when a specific tick species was encountered within the window for prophylaxis or prior to disease symptoms manifesting can help prevent prolonged or delayed medical testing and additional associated medical costs (Adrion et al. 2015). In light of recent evidence showing differences in pathogen transmission dynamics between the two main Lyme disease vectors in the United States (Couper et al. 2020), a photo-based surveillance system is also a potential avenue for rapid communication regarding risk levels encountered.

There are some limitations with this specific study as well as being inherent to this method of accurately identifying ticks from photographs. Both TickReport and TickEncounter received predominantly three tick species (*I. scapularis*, *D. variabilis*, and *A. americanum*). We recognize that by chance, identification accuracy could be falsely inflated by the increased likelihood of one of these three ticks being the correct identification. However, we used a classification accuracy analysis to mitigate this potential bias, as this method corrects for imbalance in sample proportions for determining accuracy, sensitivity, and specificity (Dziak et al. 2012). More rarely encountered ticks, or ones that appear visually similar within overlapping ranges (e.g., *D. variabilis* and *D. andersoni* in the western United States) were more likely to be incorrectly identified than ticks with more distinct ranges and appearance. Contextual parameters such as geographic location, travel history, and season in addition to a suitable size reference are still relevant to proply identifying a tick by photograph. Photos provided without these key pieces of information may limit the ability to correctly identify less commonly occurring ticks, or new imports. However, this is not only a critique of photo-based identification. Despite microscopy, *H. longicornis* was misidentified as the native *Haemaphysalis leporispalustris* for decades until a massive infestation on a New Jersey farm in 2017 (Rainey et al. 2018) brought their presence to light. This example demonstrates that visual identification alone can be mistaken if one is simply not expecting to encounter something new. A likely lesser limitation, as explained above, was the chance that the photograph identified by TickSpotters was not the same tick ultimately mailed into and identified by TickReport. Cross-referencing of matched submissions was conducted as carefully as possible (same email address within 48 h) but there may have been an insignificant number of false negatives or false positives. Finally, TickSpotters often receives photo specimens of non-tick arthropods or things that citizens believe are ticks (roughly 5% of submissions), but there were no matching submissions of non-tick specimens from TickReport to include in this analysis; therefore 'not-a-tick' identification ability could not be verified. Nevertheless, results from this study indicate that photo-based tick surveillance conducted by entomology experts is a valid and accurate method for rapid identification of commonly encountered ticks regardless of engorgement and can provide an important public health service in response to a significant and growing tickborne disease incidence.

Acknowledgments

This work was supported, in part, by the USDA National Institute of Food and Agriculture, URI Cooperative Extension, and the Rhode Island Agricultural Experiment Station. Additional support was provided by The Apple Pickers Foundation and the University of Massachusetts Laboratory of Medical Zoology. All authors declare that they have no conflicts of interest.

References Cited

Adrion, E. R., J. Aucott, K. W. Lemke, and J. P. Weiner. 2015. Health care costs, utilization and patterns of care following Lyme disease. PLoS ONE 10: e0116767.

Allen, J., S. Graziadio, and M. Power. 2017. A Shiny Tool to explore prevalence, sensitivity, and specificity on Tp, Fp, Fn, and Tn. NIHR Diagnostic Evidence Co-operative, Newcastle upon Tyne, United Kingdom.

Aronowitz, R. A. 1991. Lyme disease: the social construction of a new disease and its social consequences. Milbank Q. 69: 79–112.

Auwaerter, P. G., J. S. Bakken, R. J. Dattwyler, J. S. Dumler, J. J. Halperin, E. Mcsweegan, R. B. Nadelman, S. O. Connell, E. D. Shapiro, S. K. Sood, et al. 2011. Antiscience and ethical concerns associated with advocacy of Lyme disease. Lancet Infect. Dis. 11: 713–719.

Christenson, M., X. Lee, S. Larson, D. H. Johnson, J. Jensen, M. Meller, and S. Paskewitz. 2017. Occurrence of *Amblyomma americanum* (Acari: Ixodidae) and human infection with *Ehrlichia chaffeensis* in Wisconsin, 2008–2015. J. Med. Entomol. 54: 752–756.

Couper, L. I., Y. Yang, X. F. Yang, and A. Swei. 2020. Comparative vector competence of North American Lyme disease vectors. Parasit. Vectors 13: 29.

Dahlgren, F. S., E. J. Mandel, J. W. Krebs, R. F. Massung, and J. H. McQuiston. 2011. Increasing incidence of *Ehrlichia chaffeensis* and *Anaplasma phagocytophilum* in the United States, 2000–2007. Am. J. Trop. Med. Hyg. 85: 124–131.

Domencich, T. A., and D. McFadden. 1975. Urban travel demand - a behavioral analysis (No. Monograph).

Dziak, J. J., D. L. Coffman, S. T. Lanza, and R. Li. 2012. Sensitivity and specificity of information criteria (Methodology Center Technical Report 12-119). The Methodology Center, The Pennsylvania State University, University Park, PA.

Eisen, R. J., L. Eisen, and C. B. Beard. 2016. County-scale distribution of *Ixodes scapularis* and *Ixodes pacificus* (Acari: Ixodidae) in the continental United States. J. Med. Entomol. 53: 349–386.

Falco, R. C., T. J. Daniels, V. Vinci, D. McKenna, C. Scavarda, and G. P. Wormser. 2018. Assessment of duration of tick feeding by the scutal index reduces need for antibiotic prophylaxis after *Ixodes scapularis* tick bites. Clin. Infect. Dis. 67: 614–616.

Fogel, J., and G. S. Chawla. 2017. Susceptibility, likelihood to be diagnosed, worry and fear for contracting Lyme disease. J. Infect. Public Health 10: 64–75.

Herrington, J. E. 2004. Risk perceptions regarding ticks and Lyme disease: a national survey. Am. J. Prev. Med. 26: 135–140.

Hines, D., and S. L. Sibbald. 2015. Citizen science: exploring its application as a tool for prodromic surveillance of vector-borne disease. Can. Commun. Dis. Rep. 41: 63–67.

Johnson, T. L., K. A. Boegler, R. J. Clark, M. J. Delorey, J. K. H. Bjork, F. M. Dorr, E. K. Schiffman, D. F. Neitzel, A. J. Monaghan, and R. J. Eisen. 2018. An acarological risk model predicting the density and distribution of host-seeking *Ixodes scapularis* nymphs in Minnesota. Am. J. Trop. Med. Hyg. 98: 1671–1682.

Jordan, R. A., and A. Egizi. 2019. The growing importance of lone star ticks in a Lyme disease endemic county: passive tick surveillance in Monmouth County, NJ, 2006–2016. PLoS ONE 14: e0211778.

Koffi, J. K., J. Savage, K. Thivierge, L. R. Lindsay, C. Bouchard, Y. Pelcat, and N. H. Ogden. 2017. Evaluating the submission of digital images as a method of surveillance for *Ixodes scapularis* ticks. Parasitology 144: 877–883.

Kopsco, H. L., R. Krell, T. N. Mather, and N. P. Connally. In review. Identifying trusted sources of Lyme disease prevention information among internet users. Zoonoses and Public Health.

Landis, J. R., and G. G. Koch. 1977. The measurement of observer agreement for categorical data. Biometrics 33: 159–174.

Lee, X., D. S. Murphy, D. Hoang Johnson, and S. M. Paskewitz. 2019. Passive animal surveillance to identify ticks in Wisconsin, 2011–2017. Insects 10: 289.

Lewis, J., C. R. Boudreau, J. W. Patterson, J. Bradet-Legris, and V. K. Lloyd. 2018. Citizen science and community engagement in tick surveillance-a Canadian case study. Healthcare (Basel) 6: 22.

Mather, T. N., and M. E. Mather. 1990. Intrinsic competence of three ixodid ticks (Acari) as vectors of the Lyme disease spirochete. J. Med. Entomol. 27: 646–650.

Mather, T. N., M. C. Nicholson, E. F. Donnelly, and B. T. Matyas. 1996. Entomologic index for human risk of Lyme disease. Am. J. Epidemiol. 144: 1066–1069.

Mogg, M., H.-H. Wang, A. Baker, Z. Derouen, J. Borski, and W. E. Grant. 2020. Increased incidence of *Ehrlichia chaffeensis* infections in the United States, 2012 through 2016. Vector Borne Zoonotic Dis., online ahead of print. doi:10.1089/vbz.2019.2595.

Nelder, M. P., C. B. Russell, K. M. Clow, S. Johnson, J. S. Weese, K. Cronin, F. Ralevski, C. M. Jardine, and S. N. Patel. 2019. Occurrence and distribution of *Ambylomma americanum* as determined by passive surveillance in Ontario, Canada (1999–2016). Ticks Tick. Borne. Dis. 10: 146–155.

Nieto, N. C., W. T. Porter, J. C. Wachara, T. J. Lowrey, L. Martin, P. J. Motyka, and D. J. Salkeld. 2018. Using citizen science to describe the prevalence and distribution of tick bite and exposure to tick-borne diseases in the United States. PLoS ONE 13: e0199644.

Ogden, N. H., L. Trudel, H. Artsob, I. K. Barker, G. Beauchamp, D. F. Charron, M. A. Drebot, T. D. Galloway, R. O'Handley, R. A. Thompson, et al. 2006. *Ixodes scapularis* ticks collected by passive surveillance in Canada: analysis of geographic distribution and infection with Lyme borreliosis agent *Borrelia burgdorferi*. J. Med. Entomol. 43: 600–609.

Palmer, J. R. B., A. Oltra, F. Collantes, J. A. Delgado, J. Lucientes, S. Delacour, M. Bengoa, R. Eritja, and F. Bartumeus. 2017. Citizen science provides a reliable and scalable tool to track disease-carrying mosquitoes. Nat. Commun. 8: 916.

Pepin, K. M., R. J. Eisen, P. S. Mead, J. Piesman, D. Fish, A. G. Hoen, A. G. Barbour, S. Hamer, and M. A. Diuk-Wasser. 2012. 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. 86: 1062–1071.

Porter, W. T., P. J. Motyka, J. Wachara, Z. A. Barrand, Z. Hmood, M. McLaughlin, K. Pemberton, and N. C. Nieto. 2019. Citizen science informs human-tick exposure in the Northeastern United States. Int. J. Health Geogr. 18: 9.

Rainey, T., J. L. Occi, R. G. Robbins, and A. Egizi. 2018. Discovery of *Haemaphysalis longicornis* (Ixodida: Ixodidae) parasitizing a sheep in New Jersey, United States. J. Med. Entomol. 55: 757–759.

Rand, P. W., E. H. Lacombe, R. Dearborn, B. Cahill, S. Elias, C. B. Lubelczyk, G. A. Beckett, and R. P. Smith, Jr. 2007. Passive surveillance in Maine, an area emergent for tick-borne diseases. J. Med. Entomol. 44: 1118–1129.

Ripoche, M., S. Gasmi, A. Adam-Poupart, J. K. Koffi, L. R. Lindsay, A. Ludwig, F. Milord, N. H. Ogden, K. Thivierge, and P. A. Leighton. 2018. Passive tick surveillance provides an accurate early signal of emerging Lyme disease risk and human cases in southern Canada. J. Med. Entomol. 55: 1016–1026.

Rosenberg, R., N. P. Lindsey, M. Fischer, C. J. Gregory, A. F. Hinckley, P. S. Mead, G. Paz-Bailey, S. H. Waterman, N. A. Drexler, G. J. Kersh, et al. 2018. Vital signs: trends in reported vectorborne disease cases—United States and territories, 2004-2016. MMWR. Morb. Mortal. Wkly Rep. 67: 496–501.

Sonenshine, D. E. 2018. Range expansion of tick disease vectors in North America: implications for spread of tick-borne disease. Int. J. Environ. Res. Pub. Health 15: 478.

Springer, Y. P., L. Eisen, L. Beati, A. M. James, and R. J. Eisen. 2014. Spatial distribution of counties in the continental United States with records of occurrence of *Amblyomma americanum* (Ixodida: Ixodidae). J. Med. Entomol. 51: 342–351.

Stromdahl, E. Y., R. M. Nadolny, G. J. Hickling, S. A. Hamer, N. H. Ogden, C. Casal, G. A. Heck, J. A. Gibbons, T. F. Cremeans, and M. A. Pilgard. 2018. *Amblyomma americanum* (Acari: Ixodidae) ticks are not vectors of the Lyme disease agent, *Borrelia burgdorferi* (Spirocheatales: Spirochaetaceae): a review of the evidence. J. Med. Entomol. 55: 501–514.

Tulloch, J. S. P., L. McGinley, F. Sánchez-Vizcaíno, J. M. Medlock, and A. D. Radford. 2017. The passive surveillance of ticks using companion animal electronic health records. Epidemiol. Infect. 145: 2020–2029.

Ventola, C. L. 2015. The antibiotic resistance crisis: part 1: causes and threats. P T. 40: 277–283.

Wormser, G. P., R. J. Dattwyler, E. D. Shapiro, J. J. Halperin, A. C. Steere, M. S. Klempner, P. J. Krause, J. S. Bakken, F. Strle, G. Stanek, et al. 2006. The clinical assessment, treatment, and prevention of Lyme disease, human granulocytic anaplasmosis, and babesiosis: clinical practice guidelines by the Infectious Diseases Society of America. Clin. Infect. Dis. 43: 1089–1134.

Xu, G., T. N. Mather, C. S. Hollingsworth, and S. M. Rich. 2016. Passive surveillance of *Ixodes scapularis* (Say), their biting activity, and associated pathogens in Massachusetts. Vector Borne Zoonotic Dis. 16: 520–527.

Xu, G., P. Pearson, and S. M. Rich. 2018. *Ehrlichia muris* in *Ixodes cookei* Ticks, Northeastern United States, 2016–2017. Emerg. Infect. Dis. 24: 1143–1144.

Xu, G., P. Pearson, E. Dykstra, E. S. Andrews, and S. M. Rich. 2019. Human-biting *Ixodes* ticks and pathogen prevalence from California, Oregon, and Washington. Vector Borne Zoonotic Dis. 19: 106–114.

Yeh, M. T., J. M. Bak, R. Hu, M. C. Nicholson, C. Kelly, and T. N. Mather. 1995. Determining the duration of *Ixodes scapularis* (Acari: Ixodidae) attachment to tick-bite victims. J. Med. Entomol. 32: 853–858.