Comparison of Active and Passive Vector Surveillance Systems in Lyme Endemic Minnesota and their Correlation to Human Disease

Kathleen Angell, Janet Jarnefeld, Elizabeth K Schiffman, M Jana Broadhurst, James Dong, Abraham Degarege, Roberto Cortinas, David M Brett-Major

Submitted to: JMIR Public Health and Surveillance
on: August 02, 2024

Table of Contents

Comparison of Active and Passive Vector Surveillance Systems in Lyme Endemic Minnesota and their Correlation to Human Disease

Kathleen Angell[1] MPH; Janet Jarnefeld[2] BS; Elizabeth K Schiffman[3] MA, MPH; M Jana Broadhurst[1] PhD, MD; James Dong[1] PhD; Abraham Degarege[1] PhD; Roberto Cortinas[4] PhD, DVM; David M Brett-Major[1] MPH, MD

[1]University of Nebraska Medical Center Omaha US
[2]Metropolitan Mosquito Control District St. Paul US
[3]Minnesota Department of Health St. Paul US
[4]University of Nebraska Lincoln Lincoln US

Corresponding Author:
Kathleen Angell MPH
University of Nebraska Medical Center
984355 Nebraska Medical Center
Omaha
US

Abstract

Background: Vector surveillance is often used to predict tickborne diseases in endemic regions. Active and passive vector surveillance systems offer differing benefits and limitations, understanding how the outputs of these systems differ and how they correlate to human disease is essential to public health decision making. Active and passive vector surveillance systems in place in Minnesota between 2018 and 2023 present an opportunity for comparison between these surveillance methods.

Objective: To (i) analyze, compare, and contrast the results of active and passive vector surveillance programs; and (ii) explore how well these sources predict human risk of Lyme disease.

Methods: Descriptive statistics were performed to evaluate characteristics of each surveillance method with chi square tests or repeated-ANOVA to assess differences in seasonality, life stage, and genus of ticks between different datasets. Correlation to human cases of Lyme disease was analyzed using negative binomial regression models.

Results: There are differences between the data sources in tick life stage and genus proportions as well as seasonality of tick rates. Active surveillance conducted by the Metropolitan Mosquito Control District (MMCD) using small mammal trapping had a majority of larval I. scapularis ticks. In contrast, passive surveillance by iNaturalist had a majority of adult D. variabilis ticks. Observations in both data sources were skewed to the early third of the tick season, although this was more exaggerated in iNaturalist data. Observations of ticks from both data sources positively correlated with human cases of Lyme disease.

Conclusions: Observed differences in tick characteristics between the two data sources may represent real differences between tick populations and human encounters with them. Some differences may be explained by observation, reporting, and sampling biases. Increased observations of ticks at the beginning of the season indicates potential utility of enhanced human Lyme disease surveillance at that time and may apply to other tick-borne disease risk management. These One Health findings signal an opportunity for early identification of high tickborne disease years through integrated active and passive tick surveillance that informs the conduct of human disease surveillance.

(JMIR Preprints 02/08/2024:65034)
DOI: https://doi.org/10.2196/preprints.65034

2) If accepted for publication in a JMIR journal, would you like the PDF to be visible to the public?

✓ **Yes, please make my accepted manuscript PDF available to anyone at any time (Recommended).**

Yes, but please make my accepted manuscript PDF available only to logged-in users; I understand that the title and abstract will remain v

Yes, but only make the title and abstract visible (see Important note, above). I understand that if I later pay to participate in <a href="http

Original Manuscript

Original Paper

Kathleen E Angell*[a], Janet Jarnefeld[b], Elizabeth K Schiffman[c], M Jana Broadhurst[d,e], James Dong[f], Abraham Degarege[a], Roberto Cortinas[g], David M Brett-Major[a,d]

- **a.** Department of Epidemiology, College of Public Health, University of Nebraska Medical Center, Omaha, NE
- **b.** Metropolitan Mosquito Control District, Saint Paul, MN
- **c.** Minnesota Department of Health, St. Paul, MN
- **d.** Global Center for Health Security, University of Nebraska Medical Center, Omaha, NE
- **e.** Department of Pathology, Microbiology, and Immunology, College of Medicine, University of Nebraska Medical Center, Omaha, NE
- **f.** Department of Biostatistics, College of Public Health, University of Nebraska Medical Center, Omaha, NE
- **g.** School of Veterinary Medicine and Biomedical Sciences, University of Nebraska, Lincoln, NE

***Corresponding Author:** Kathleen E. Angell, Kathleen.Angell@UNMC.edu
Phone: 612-239-6140, Office: 984395 Nebraska Medical Center, Omaha, NE 68198-4395

Comparison of Active and Passive Vector Surveillance Systems in Lyme Endemic Minnesota and their Correlation to Human Disease

Abstract

Background: Vector surveillance is often used to predict tickborne diseases in endemic regions. Active and passive vector surveillance systems offer differing benefits and limitations, understanding how the outputs of these systems differ and how they correlate to human disease is essential to public health decision making. Active and passive vector surveillance systems in place in Minnesota between 2018 and 2023 present an opportunity for comparison between these surveillance methods.

Objective: To (i) analyze, compare, and contrast the results of active and passive vector surveillance programs; and (ii) explore how well these sources predict human risk of Lyme disease.

Methods: Descriptive statistics were performed to evaluate characteristics of each surveillance method with chi square tests or repeated-ANOVA to assess differences in seasonality, life stage, and genus of ticks between different datasets. Correlation to human cases of Lyme disease was analyzed using negative binomial regression models.

Results: There are differences between the data sources in tick life stage and genus proportions as well as seasonality of tick rates. Active surveillance conducted by the Metropolitan Mosquito Control District (MMCD) using small mammal trapping had a majority of larval *I. scapularis* ticks. In contrast, passive surveillance by iNaturalist had a majority of adult *D. variabilis* ticks. Observations in both data sources were skewed to the early third of the tick season, although this was more exaggerated in iNaturalist data. Observations of ticks from both data sources positively correlated with human cases of Lyme disease.

Conclusions: Observed differences in tick characteristics between the two data sources may represent real differences between tick populations and human encounters with them. Some differences may be explained by observation, reporting, and sampling biases. Increased observations of ticks at the beginning of the season indicates potential utility of enhanced human Lyme disease surveillance at that time and may apply to other tick-borne disease risk management. These One Health findings signal an opportunity for early identification of high tickborne disease years through integrated active and passive tick surveillance that informs the conduct of human disease surveillance.

Keywords

Ixodes Scapularis Tick; Dermacentor Variabilis Tick; Vector; Disease Surveillance; Lyme disease; Midwest; Crowdsourcing

Introduction

Vector surveillance is an important public health measure to predict the risk of tickborne diseases in a population [1, 2]. Lyme disease, a nationally notifiable disease, is the most common tickborne disease reported in Minnesota and the United States [3]. In the upper Midwest, *Borrelia burgdorferi*, the causative agent for Lyme disease is transmitted to humans by the *Ixodes scapularis* tick which can also transmit the pathogens that cause anaplasmosis, babesiosis, and ehrlichiosis [3]. *Dermacentor variabilis*, another commonly encountered tick in the region, can spread rocky mountain spotted fever and tularemia, emerging diseases of concern in Minnesota [3, 4]. Given the threat of these diseases, it is important to assess vector surveillance methods available for use in the state.

There are several forms of active and passive surveillance systems in place in Minnesota. Active vector surveillance occurs through systematic field collection using methods such as flag sampling, walk sampling, removing ticks from hosts, or CO_2 trapping [5]. The Metropolitan Mosquito Control District (MMCD) [6] assesses tick abundance through removal of ticks from trapped small mammals in the seven-county Minneapolis/St. Paul metropolitan region. In contrast, passive vector surveillance occurs through people voluntarily submitting ticks to or recording observations of ticks for researchers or government agencies [5, 7]. Crowd-sourced observations recorded by “citizen scientists” to the publicly available iNaturalist database represent a growing data stream in the region as well as worldwide.

In limited time span studies, combining results from crowd-sourced passive vector surveillance systems with those from active vector surveillance efforts has proved a valuable source of information in understanding the geographic spread of tick genera and helps to identify locations that may benefit from active vector surveillance [5, 8]. A longer-term, multi-year exploration of the intersection between active and passive vector surveillance datasets is needed to fully understand how they can be utilized to predict human risk of tickborne disease and inform public health decision making.

MMCD’s active vector surveillance program and iNaturalist’s crowd-sourced passive surveillance database have multiple years of overlapping data, which provides a unique opportunity to explore how their observations intersect over time. The objectives of this assessment were to (i) analyze, compare, and contrast the results of active and passive vector surveillance programs; and (ii) explore how well these sources predict human risk of Lyme disease.

Methods

Data Sources & Sets

Data used in this study were obtained from Metropolitan Mosquito Control District (MMCD), iNaturalist and the Minnesota Department of Health (MDH).

Metropolitan Mosquito Control District

MMCD began active vector surveillance in the seven-county Twin Cities metropolitan area in 1990. MMCD conducts vector surveillance through small mammal trapping from the last week of April to the last week of October across one hundred randomly selected sites using a square-mile sampling region within each township. Trapping and removal of on-host ticks is performed on a subset of sites, rotating weekly so that each location is measured three times within the twenty-seven-week sampling period.

MMCD provided tick prevalence data from 2014 to 2023 in the form of a line list with each line representing one trapped small mammal. Host species, sex, number of ticks, week of surveillance period, date, and trapping location were included. The number of *Dermacentor variabilis* and *Ixodes scapularis* ticks found on each animal was recorded stratified by life stage (larva or nymph). The number of ticks outside of those genus and life stage was recorded as other. An adult life stage category was added for each of the two genera for consistency between datasets.

iNaturalist

iNaturalist, an application founded in 2008, is used globally by citizen scientists for nature observation and identification. People anywhere can use their smartphone to collect images of flora and fauna observations and provide geographic coordinates of the observations. This open-source database represents a passive form of surveillance searchable by location, time, and genus. Verifiable observations of *Ixodidae* (hardbacked ticks) from iNaturalist for the seven-county Twin Cities metropolitan area from 2014 to 2023 were downloaded, selecting only variables of interest to the study. Verifiable observations on the iNaturalist database are those that contain valid date, location, photo, and is not of a captive/cultivated organism [9]. Each line of the dataset represented a single observation, including an identification number, date observed, geocoded location and county of observation, genus, and life stage of the tick. We imputed the week of observation to match that of the MMCD dataset based on the observation date. We categorized ticks from the iNaturalist dataset based on two main genera of interest (*D. variabilis* and *I. scapularis*) and three life stages to match the categories in the MMCD dataset. The cleaned iNaturalist dataset can be found on Harvard Dataverse [10].

Minnesota Department of Health

The MDH conducts passive surveillance of human cases of Lyme disease statewide in accordance with nationally notifiable disease requirements by the Centers for Disease Control and Prevention (CDC) [11]. The number of cases by county for the seven-county Twin Cities metropolitan area from 2014 to 2023 was provided by MDH.

Statistical Analysis

To better describe the seasonality of tick abundance, we categorized week of collection/observation as early, mid, or late tick season by dividing the twenty-seven-week MMCD collection period into thirds. Observations from iNaturalist before the beginning of the MMCD collection period were categorized as early, and from after the MMCD collection period as late. We consolidated data from both vector surveillance sources by week, with overall counts for each of the seven tick categories (*D. variabilis* larva, nymph, and adult, *I. scapularis* larva, nymph, and adult, and all other genera regardless of life stage). Total number of ticks, total number by genus and total number by life stage were calculated. Two other datasets were created, one collapsed by season and the other by year. We combined the MDH dataset containing human cases of Lyme disease with the tick dataset collapsed by year to test for a correlation between tick abundance and human cases of tick-borne disease.

We performed descriptive statistics and used multivariable statistical models to understand seasonal and spatial distribution, life stage and genera of recorded ticks and their relationship to human case rates. Chi-square tests or repeated-ANOVA were used to assess differences in seasonality, life stage, and genus of ticks between different datasets. Correlation matrices were produced to examine the association between human cases of Lyme disease and the total number of ticks unadjusted, categorized by genus and life stage, both in the combined dataset and stratified by data source (MMCD and iNaturalist). Generalized estimating equation (GEE) models with negative binomial distributions were conducted with the number of human cases as the outcome variable and various tick variables as the main predictors (overall total ticks, total *I. scapularis*, total *D. variabilis*, total larva, total nymph, and total adult ticks), adjusted for county. These six GEE models were conducted on each of the three datasets (MMCD, iNaturalist, and combined datasets) for a total of eighteen models. The datasets utilized the maximum number of years data available as illustrated in Figure 4 with the combined and iNaturalist sets using years 2018 through 2023, and the MMCD using 2014 through 2023. Significance was set at an alpha of 0.05. Data was analyzed using SAS in version 9.4.

The decision to explore tick prevalence rather than density or intensity in the MMCD dataset was made to maintain consistency with the iNaturalist dataset after a review of vector surveillance literature revealed prevalence, density and intensity to be equally appropriate measures, each with their own sets of strengths and limitations [4, 12-15]. We performed a sensitivity analysis to explore the effects of employing tick prevalence versus density in the MMCD dataset. Density was calculated by dividing the number of ticks by the number of small mammals trapped, including mammals with no attached ticks.

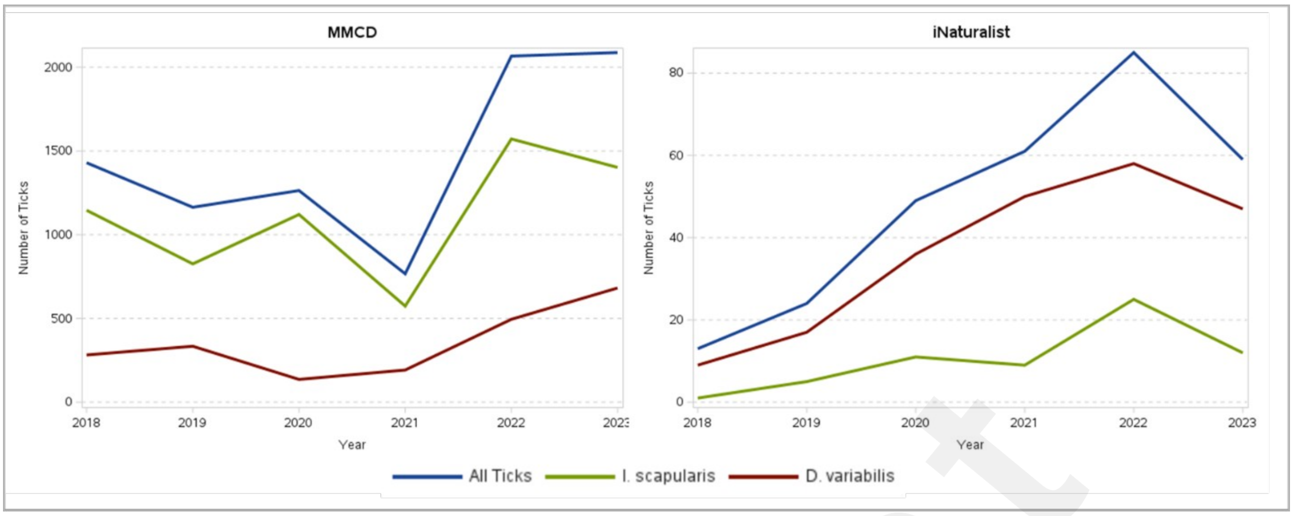
Results

There were 9,071 ticks observed in the two datasets between 2018 and 2023 with 97% (8,870 ticks) of observations from MMCD and 3.2% (291 ticks) from the iNaturalist database (Table 1). Twenty-four, or 8.2% of iNaturalist observations took place before the beginning of the MMCD collection period and nine after the end of the period. The decision to restrict the datasets to 2018 to 2023 in the analysis was made due to low numbers in the iNaturalist dataset in the earlier years of its founding. The growing popularity of the iNaturalist application makes year to year comparison difficult as a steady rise is seen each year until 2021 (Figure 1). Among MMCD observations, tick abundance has characteristic highs and lows from one year to another, though low tick numbers were seen in 2021 in concert with a major drought.

Table 1. Vector Characteristics by Data Source (2018-2023)

	MMCD	**iNaturalist**	**Total**
	N (%)		
All Ticks	8780 (97)	291 (3.2)	9071 (100)
Genera			
Ixodes	6641 (76)	63 (22)	6704 (74)
Dermacentor	2121 (24)	217 (75)	2338 (26)
Other	18 (0.2)	11 (3.8)	29 (0.3)
Season			
Early	3706 (42)	225 (77)	3931 (43)
Mid	4065 (46)	43 (15)	4108 (45)
Late	1009 (12)	23 (7.9)	1032 (11)
Life stage			
Larvae	7447 (85)	0 (0)	7447 (82)
Nymph	1315 (15)	11 (3.9)	1326 (15)
Adult	0 (0)	269 (96)	269 (3.0)
County			
Anoka	3151 (36)	64 (22)	3215 (35)
Carver	161 (1.8)	8 (2.7)	169 (1.9)
Dakota	1041 (12)	68 (23)	1109 (12)
Hennepin	927 (11)	45 (16)	972 (11)
Ramsey	501 (5.7)	34 (12)	538 (5.9)
Scott	389 (4.4)	10 (3.4)	399 (4.4)
Washington	2610 (30)	62 (21)	2672 (30)

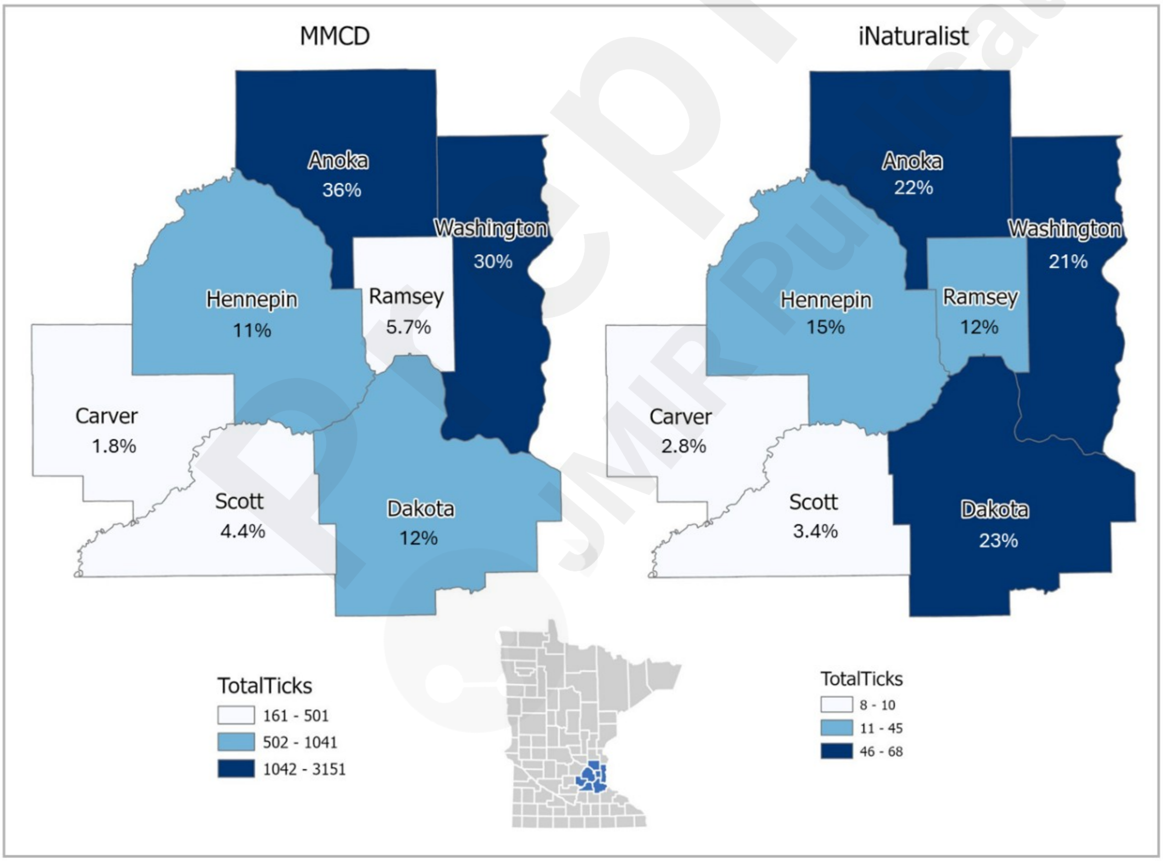
Figure 1. Ticks by Data Source, Year and Genus



Spatial Distribution

In the combined dataset, the majority (65%) of ticks regardless of genus and life stage were seen in Washington and Anoka counties, driven by trends in the MMCD dataset (Figure 2). The iNaturalist dataset shows a more even distribution across the metropolitan region with the exceptions of Carver and Scott counties which only represent 2.8% and 3.4% of observed ticks, respectively.

Figure 2. Heat Map of Ticks by Data Source and County



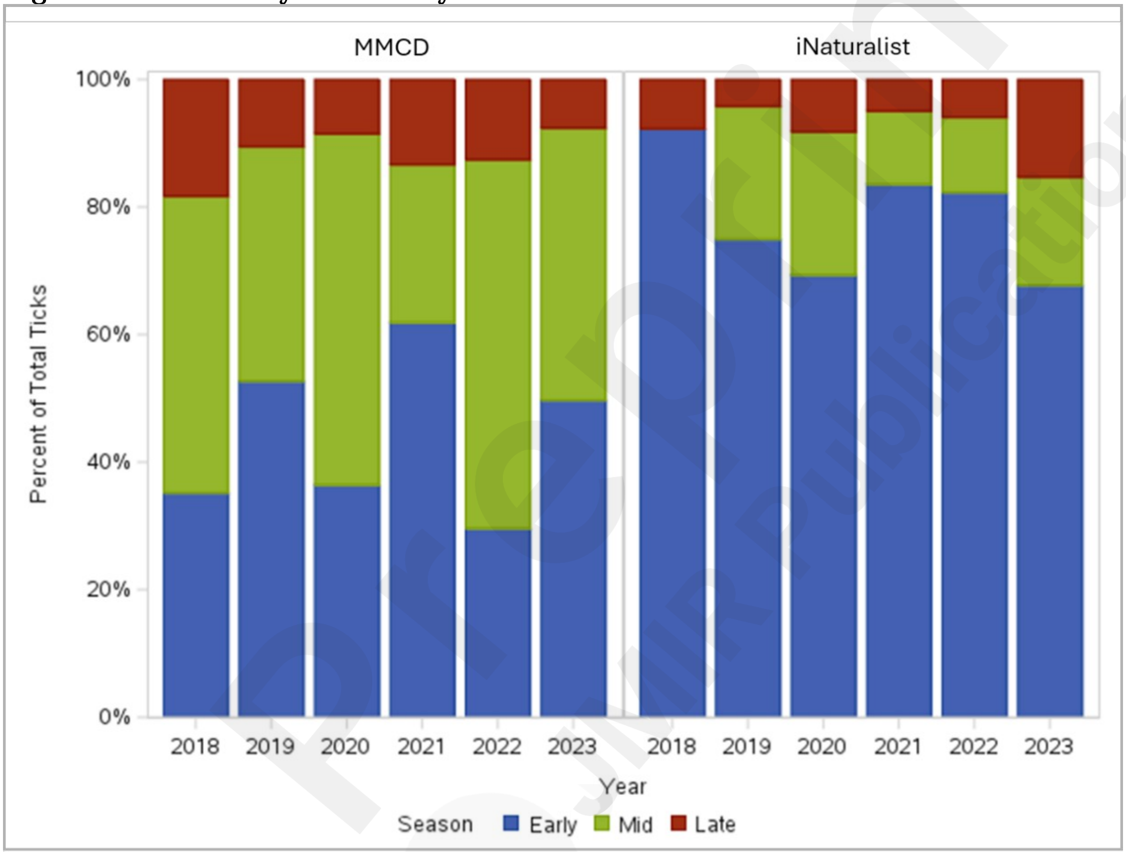
Genus Proportion

Seventy-six percent of tick observations in the MMCD dataset were *I. scapularis* compared to 24% in iNaturalist. The majority (78%) of iNaturalist observations were *D. variabilis* compared to 22% of MMCD observations (Table 1), a significant difference using the chi-square test (p<0.0001).

Seasonal Distribution

Seventy-seven percent of tick observations in the iNaturalist dataset were observed in the early part of the tick season (last week of April to the second to last week in June) with the remaining observations waning throughout the mid and late parts of the season (last week in June through the last week in October). In contrast, MMCD observations are evenly spread across the early and middle of the season with only 12% of observations in the late season. This trend is consistent from year to year (Figure 3). This difference in seasonality between the datasets reaches statistical significance in a chi-square test with p<0.0001. When this analysis is stratified by tick genera, this same difference in seasonality was seen among *I. scapularis* ticks with 70%, 0% and 30% of iNaturalist observations taking place in early, mid, and late season respectively, compared to 30%, 56% and 14% for the MMCD dataset. In contrast, among *D. variabilis* ticks, matching trends in seasonality are seen in both datasets with most observations (80%) in the early part of the season, and the remaining waning in tandem throughout the rest of the season.

Figure 3. Seasonality of Ticks by Data Source and Year



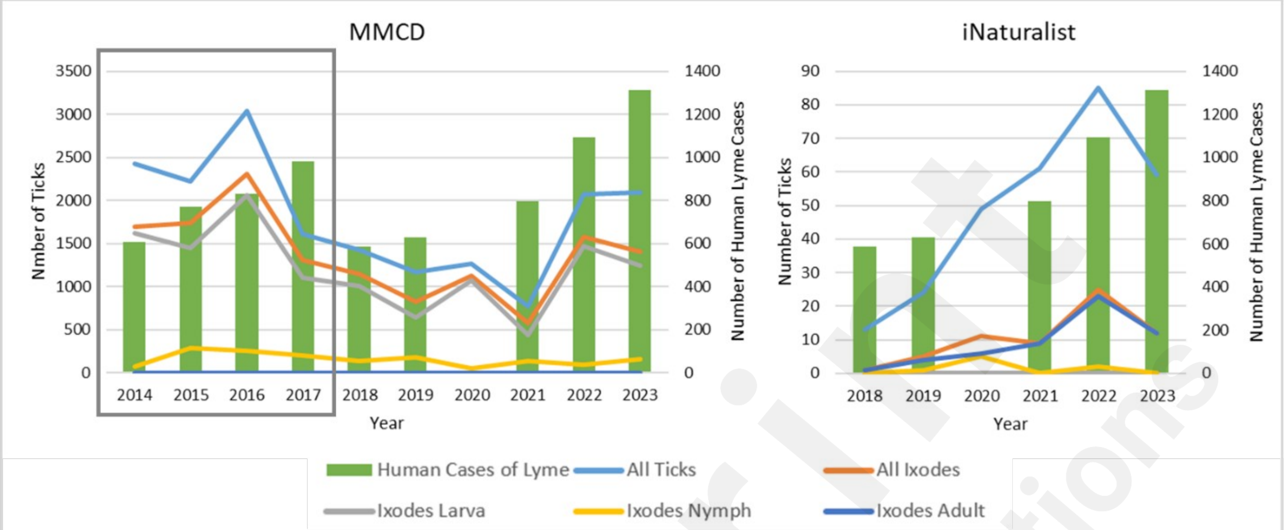
Life Stage Proportion

Ninety-six percent of ticks in the iNaturalist dataset are in the adult life stage, 3.9% in the nymphal life stage, and no larva. The MMCD dataset has 86% of ticks in the larval life stage, 14% in the nymphal, and no adult ticks. In a chi square analysis, this difference in life stage proportions reaches statistical significance with p<0.0001.

Association with Human Cases of Lyme Disease

Data provided from MDH shows an overall rise in Lyme disease cases since 2018 (Figure 4). Surveillance data was not available in 2020 leading to missing information for that year. Pearson Correlation Coefficients were generated to assess overall association between human cases of Lyme disease and ticks in the joint dataset as well as stratified by data source. The combined dataset shows low degree correlation between human cases and all tick categories with the highest coefficients for all adult ticks (r(66)= .28, p= .024) and adult *I. scapularis* ticks (r(66)= .29, p= .020). The MMCD dataset also shows low or moderate degree correlation among all tick categories, although fail to meet significance at an alpha of 0.05 except total *D. variabilis* ticks (r(35)= .36, p= .036). In the iNaturalist dataset, human cases have a low to moderate degree correlation with all tick categories, the most notable being total ticks (r(31)= .46, p= .009), total number of *I. scapularis* (r(31)= .45, p= .012) and *I. scapularis* in the adult life stage (r(31)= .45, p= .012).

Figure 4. Tick Counts Compared to Human Cases of Lyme Disease by Data Source and Year



▭ - Additional years of tick data available with the MMCD dataset; these additional years were included in the MMCD stratified negative binomial models.

*Human cases of Lyme Borreliosis missing for 2020 due to lapse in surveillance

Negative binomial regression adjusting for county was performed on the combined dataset as well as stratified by data source (MMCD and iNaturalist) (Supplementary Table 1). All three datasets found significant positive correlation to human cases of Lyme, that is, as number of ticks increases, so do human cases. The combined dataset has strongest association between all ticks in the adult life stage and *D. variabilis* ticks (*b*=0.018, 95% CI = 0.012, 0.024, p<.0001; *b*=0.001, 95% CI = 0.0007, 0.002, p<.0001, respectively). Positive adjusted association between all tested variables and human cases of Lyme is also seen in the MMCD dataset with all nymphal stage ticks posing the strongest association (*b*=0.005, 95% CI = 0.003, 0.008, p<.0001). Similar positive association is seen in the iNaturalist dataset with all *I. scapularis* ticks and all nymphal stage ticks leading the association (*b*=0.135, 95% CI = 0.085, 0.184, p<.0001; *b*=0.114, 95% CI = 0.046, 0.182, p= 0.001 respectively).

Sensitivity Analysis

A sensitivity analysis to explore the effects of employing vector density rather than prevalence to represent MMCD active vector surveillance results did not substantially change the findings. Ramsey County transitioned from low to high burden in the spatial distribution. Association to human cases did not differ in directionality or statistical significance.

Discussion

The MMCD dataset with its randomly chosen sites using a square mile sampling region provides a baseline in tick abundance for a given region regardless of human population size or likeliness of outdoor human activities such as hiking. *I.* scapularis ticks have a well-known preference for deciduous forests [4], which is reflected in the spatial distribution in the MMCD dataset with counties in the deciduous northeast region of the metropolitan area (Anoka and Washington) contributing to the highest number of ticks, and counties in the southwest area (Carver and Scott) contributing the least. The iNaturalist dataset has a more even distribution of tick observations across the region apart from Carver and Scott counties which collectively account for 6% of tick abundance; in addition to the lack of deciduous forests, these are also the smallest counties by population in the area which could result in fewer observations in a passive dataset like iNaturalist [16]. However, the next smallest counties by population are Washington, Anoka, and Dakota which represent the largest contributors for tick observations (Table 2). Considering baseline tick abundance determined from MMCD data and the populations for the included counties, iNaturalist data appears to reflect where people are encountering ticks, and not necessarily population sizes.

Table 2. Minnesota Population[1] & Reported Lyme Disease Cases[2] by County

County	Population	Lyme Cases	Lyme Case Rate[3]
Hennepin	1,293,651	1,358	105
Ramsey	551,195	791	144
Dakota	444,985	568	128
Anoka	368,280	687	187
Washington	274,589	767	279
Scott	154,395	147	95
Carver	110,136	100	91

1 - According to the 2022 Minnesota State Demographic Center Survey[7]
2 - 2018-2023 (without 2020 due to lapse in surveillance)
3 - Number of cases per 100,000 persons from 2018 to 2023

There is a marked difference in genus proportions observed between the two datasets with *I. scapularis* ticks dominating the MMCD dataset and *D. variabilis* ticks leading the iNaturalist dataset. MMCD's vector monitoring program was originally initiated to explore the spread of *I. scapularis* ticks in the region and has successfully focused on those observations. Observations in the iNaturalist dataset, on the other hand, reflect human encounters with ticks through natural exposure which can result in selection biases discussed in Table 3.

The datasets reveal differences in when tick observations are happening within the tick season (late April through late October). MMCD's surveillance method, using small mammal trapping, means that they are less likely to observe adult ticks which prefer larger mammals for feeding and are more likely to emerge late in the season; this could be one explanation for the lack of late observations among this dataset. The iNaturalist dataset, however, also sees less observations at the end of the tick season with an even greater skew to the first third of the season despite a predisposition for adult life stage tick observations. This could be due to multiple biasing factors discussed in Table 3 including observations preceding the beginning of the MMCD collection period; waning interest in the application over the course of the season; or decreased time spent outdoors over the season. Minnesotans may be especially eager to spend time outdoors when spring comes after a long winter. As the temperatures rise and become less habitable in the peak of summer, they may be more likely to retreat again into shaded and air-conditioned comfort. Additionally, 8% of iNaturalist observations take place before the MMCD collection period starts which suggests a potential need to expand the active vector surveillance period to better capture the full season of exposure to ticks.

Opposite trends in life stage observations between the datasets are unsurprising given that early life stages generally prefer small mammals for their blood meal (MMCD), where ticks in the adult stage generally prefer larger mammals such as humans, deer, or coyotes and so are present in more observable locations (iNaturalist) [17]. Nymphal stage ticks also will commonly feed on humans which is not reflected as clearly in the data. It is possible that the lack of nymphal ticks in the iNaturalist dataset is due to their small size, consistent with the thinking that nymphs are particularly problematic for disease transmission. Nymphs may go unnoticed for long periods of time, allowing for transmission to occur rather than be interrupted by tick removal in the first 24 or so hours [18].

The correlation between ticks and human cases of Lyme disease that can be observed in figure 4 is reflected in the Pearson correlation analysis and negative binomial models exploring this relationship. Growing popularity in the iNaturalist dataset over time leads to a steady increase in overall observations of ticks from year to year; while this correlates well with the steady rise in human cases in the same timeframe, it is important to acknowledge the bias presented by growth in use of the application. Trends in rates of *I. scapularis* over time show peaks and valleys more expected of tick abundance rather than a steady rise suggesting this genus was less susceptible to the effects of the growing popularity of the application over time, perhaps due to the perceived significance of this tick among Minnesotans. It is more relevant therefore to explore the relationship between *I. scapularis* ticks and cases of Lyme; here a positive significant association with total *I. scapularis* and *I. scapularis* in the adult stage and human cases is seen. The association fails to meet significance among *I. scapularis* nymphs, likely due to low numbers of observations within this category. Regardless of the iNaturalist dataset's affinity towards observations of *D. variabilis*, it remains a valuable tool for predicting risk of tickborne disease associated with *I. scapularis* ticks.

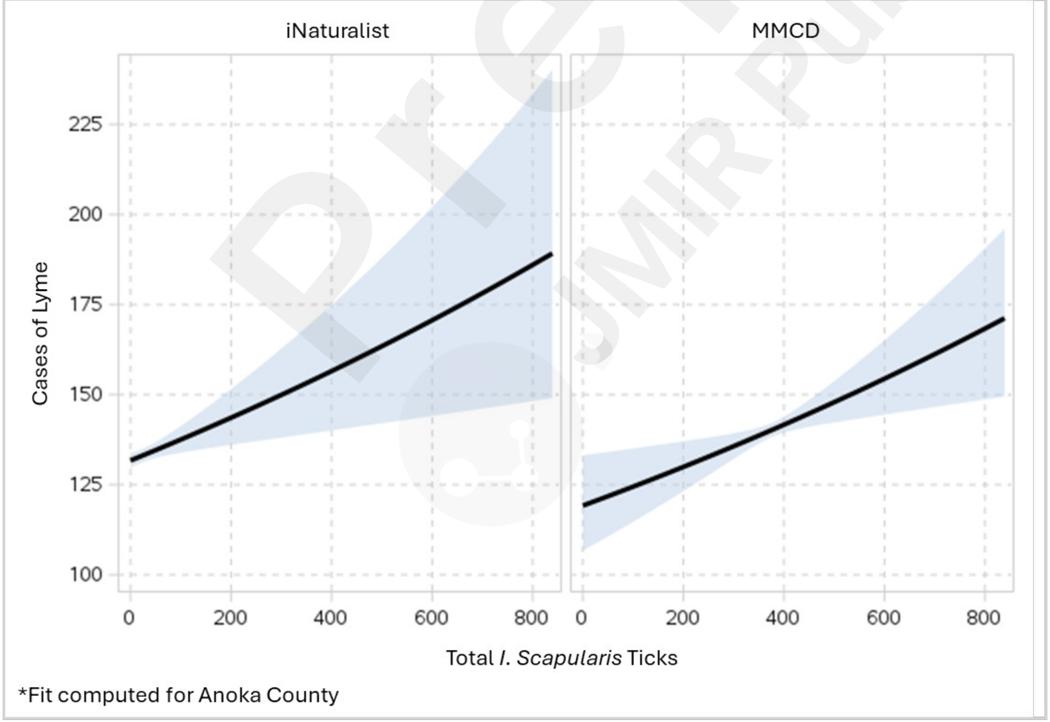
Table 3: Potential Sources of Bias in Applying Data to Human Lyme Risk Assessment

	MMCD	**iNaturalist**
Who	LOW Systematic collection methods mean ticks are recorded regardless of genus or life stage.	UNKNOWN People may be more likely to post tick observations that they have high confidence or deem significant in some way such as rare in genus or life stage, or of clinical significance.
What	HIGH Method of collection (small mammal trapping) was specifically chosen to target *I. scapularis* ticks, therefore, other genus of ticks prominent in the region may be lacking.	LOW Method of collection does not purposefully target or exclude any genus or life stage of tick.
	HIGH Small mammal trapping is more likely to result in larva or nymph ticks with few to no adults represented.	HIGH Passively collecting ticks off humans (and sometimes pets) is more likely to result in adult or nymph ticks with few to no larva represented.
	LOW Each tick is inspected under magnification by an expert to determine genus and life stage.	LOW Citizen scientists enter into the database their assessment for genus and life stage and other citizen scientists are able to vote agreement or disagreement with the assessment. We conducted a review of all observations and found very little evidence of misclassification.
Where	LOW 100 random collection sites using square mile sampling region means a consistent view of tick density over space, especially when aggregated by county.	MODERATE Collection sites are based on where users of the application live or visit which may result in hotspots in highly populated areas, public parks or along walking paths; this may provide a disjointed view of tick density at a geocoded level; however this is smoothed by looking at the data on a county level.

	MODERATE	MODERATE
	Random sampling of sites for tick collection were selected based on suitability of tick habitats rather than where people are most likely to encounter ticks, this may result in limitations for estimations of people-based encounters of the species.	People who use the application may be more likely to post more often, thus increasing density of observations in an area. We conducted a review of user repetition and found about 30 users with repeated observations which varied over both time and space. The average posting rate for a user was 1.7 (standard deviation: 2.6).
When	LOW	UNKNOWN
	Systematic collection methods mean consistent collection practice across the season. Although lack of adult ticks may bias the dataset to early tick season.	People may be more likely to post tick observations at the beginning of the tick season when encountering them is still novel.

The MMCD dataset sees significant positive associations between all measures of tick counts and human cases of Lyme. The targeted efforts of their surveillance program to gain a baseline understanding of the *I. scapularis* genus using a random selection method for surveillance sites rather than targeted to where humans interact most with tick abundant environments does not seem to make a difference when exploring association with human disease. Although the negative binomial regression estimates in all models seem low, and Pearson correlation coefficients range from low to moderate degree association, the impact of tick prevalence on human cases is considerable. Figure 5 shows predicted human cases (and 95% confidence band) of Lyme against total *I. scapularis* ticks detected from either MMCD or iNaturalist surveillance mechanisms as a depiction of the negative binomial regression results.

Figure 5. Predicted Cases of Lyme Disease with 95% Confidence Limits from Total *I. scapularis* Ticks Using Negative Binomial Regression Outputs



When the relative trends of *I. scapularis* ticks from year to year are compared between the two datasets in figure 4, we see the trends mirrored but more exaggerated in the MMCD dataset. This exaggeration is particularly noticeable in 2021 when a drought decimated the tick population; this is reflected by a steep decline that year in the MMCD dataset compared to a slight dip in the iNaturalist dataset. This could be due to larval and nymphal stage ticks more prominent in the MMCD dataset being more susceptible to drought but may also suggests that fewer ticks in the environment does not necessarily mean less human interaction with ticks and therefore cases of Lyme which continued to climb that year.

The sensitivity analysis to explore vector density rather than prevalence in the MMCD dataset did not show important differences in outputs when employing one method versus the other. Regardless of these findings, it remains important to be mindful of the context, implications and use of these methods and perform sensitivity analyses to determine the extent of any differences.

One major strength of this study is the comparison of active and passive vector surveillance systems in overlapping geographic areas over the span of multiple years. The dataset provided by MMCD represents a rare example of unbroken multi-decade active vector surveillance which contributes to our knowledge of tick abundance and habits in the region. The iNaturalist dataset is utilized globally, but seldom combined with traditional surveillance systems to fill gaps in knowledge. As with any data collection effort, both datasets are subject to bias which is discussed in Table 3. While the iNaturalist dataset presents opportunities for strengthening more traditional active surveillance methods, it also presents some weaknesses such as increased popularity of the application over time leading to difficulties in abundance comparisons from one year to the next. We feel these weaknesses and biases are outweighed by the information gathered on seasonality, genera and life stage trends in addition to its representation of people and their natural exposure to tick abundant environments. Cases of Lyme disease from MDH are recorded by county of residence which may not reflect exposure. Additionally, the case definition for reporting of Lyme disease changed in 2022, collapsing probable and confirmed cases into one probable case category; combining the probable and confirmed cases in years prior to 2022 increases the comparability from before to after the definition change, nonetheless, this represents a potential bias in the data. Additional years and expanded geographic range would be needed to adequately assess for time lag and threshold effects in the relationship between tick abundance and human cases.

Conclusions

Vector surveillance is important for understanding human risk of tick-borne diseases and tracking geographic spread of high concern genera such as *I. scapularis* which can carry pathogens that cause Lyme disease, the most common vector-borne disease in the United States. Active and passive vector surveillance systems present their own strengths and weaknesses; together, they can provide a more comprehensive understanding of tick populations, their geographic spread and their effects on human health. We have shown here that there are differences in tick genera, life stage and seasonality between the active surveillance represented by MMCD efforts, and passive surveillance represented by the iNaturalist database. The MMCD data source excels in gaining a baseline population abundance of *I. scapularis* ticks and their geographic movement over time while still serving as an accurate predictor of human disease. The iNaturalist dataset provides a view of what genera, life stage and season humans are coming into contact with ticks and reflects annual changes in human cases of Lyme. The abundance of ticks seen in the iNaturalist data source prior to the initiation of MMCD seasonal surveillance presents evidence for the expansion of MMCD surveillance period.

Acknowledgments

The authors are grateful for participation by the Metropolitan Mosquito Control District (MMCD) and Minnesota Department of Health (MDH). The authors have no conflicts of interest related to this work. These views are those of the authors and do not necessarily reflect those of any agency.

Conflicts of Interest

None declared.

Abbreviations

CDC: Centers for Disease Control and Prevention
GEE: generalized estimating equation
MDH: Minnesota Department of Health
MMCD: Metropolitan Mosquito Control District

References

1. Burtis, J.C., et al., *Predicting distributions of blacklegged ticks (Ixodes scapularis), Lyme disease spirochetes (Borrelia burgdorferi sensu stricto) and human Lyme disease cases in the eastern United States.* Ticks and tick-borne diseases, 2022. **13**(5): p. 102000.
2. Eisen, R.J. and C.D. Paddock, *Tick and Tickborne Pathogen Surveillance as a Public Health Tool in the United States.* J Med Entomol, 2021. **58**(4): p. 1490-1502.
3. *Diseases that can be transmitted by ticks. Diseases that can be Transmitted by Ticks - MN Dept of Health*. January 30, 2024; Available from: https://www.health.state.mn.us/diseases/tickborne/diseases.html.
4. Cassens, J., et al., *Environmental Drivers of Immature Ixodes scapularis in Minnesota's Metro Area.* EcoHealth, 2023. **20**(3): p. 273-285.
5. Lyons, L.A., et al., *Comparing Contributions of Passive and Active Tick Collection Methods to Determine Establishment of Ticks of Public Health Concern Within Illinois.* J Med Entomol, 2021. **58**(4): p. 1849-1864.
6. *Home: Metropolitan Mosquito Control District. MMCD*. Available from: https://mmcd.org/.
7. Maxwell, S.P., et al., *Improving Surveillance of Human Tick-Borne Disease Risks: Spatial Analysis Using Multimodal Databases.* JMIR Public Health Surveill, 2023. **9**: p. e43790.
8. Holcomb, K.M., et al., *Comparison of acarological risk metrics derived from active and passive surveillance and their concordance with tick-borne disease incidence.* Ticks Tick Borne Dis, 2023. **14**(6): p. 102243.
9. *What is a "verifiable observation" and how does it reach "research grade"? Frequently Asked Questions*. 2019, August 6; Available from: https://www.inaturalist.org/posts/26549-what-is-a-verifiable-observation-and-how-does-it-reach-research-grade.
10. *iNaturalist Ticks by Year and County (Minnesota, 2018-2023*. 2024; Available from: https://mcas-proxyweb.mcas.ms/certificate-checker?login=false&originalUrl=https%3A%2F%2Fdoi.org.mcas.ms%2F10.7910%2FDVN%2F7MCPCZ%3FMcasTsid%3D20893&McasCSRF=45effb7807cf464983ad6b582f5dc59a2fc744fb799306ab79ad02498833a865.
11. *Vaccine preventable diseases surveillance manual: Enhancing surveillance*. Manual for the Surveillance of Vaccine-Preventable Diseases 2017, November 17.
12. Bouchard, C., et al., *Associations between Ixodes scapularis ticks and small mammal hosts in a newly endemic zone in southeastern Canada: implications for Borrelia burgdorferi transmission.* Ticks Tick Borne Dis, 2011. **2**(4): p. 183-90.
13. Elias, S.P., et al., *Emergence of Ixodes scapularis (Acari: Ixodidae) in a Small Mammal Population in a Coastal Oak-Pine Forest, Maine, USA.* J Med Entomol, 2022. **59**(2): p. 725-740.
14. Ferreira, F.C., et al., *Ticks (Acari: Ixodida) on synanthropic small and medium-sized mammals in areas of the northeastern United States infested with the Asian longhorned tick, Haemaphysalis longicornis.* Int J Parasitol, 2023. **53**(14): p. 809-819.
15. White, S.A., et al., *Surveys for ticks on wildlife hosts and in the*

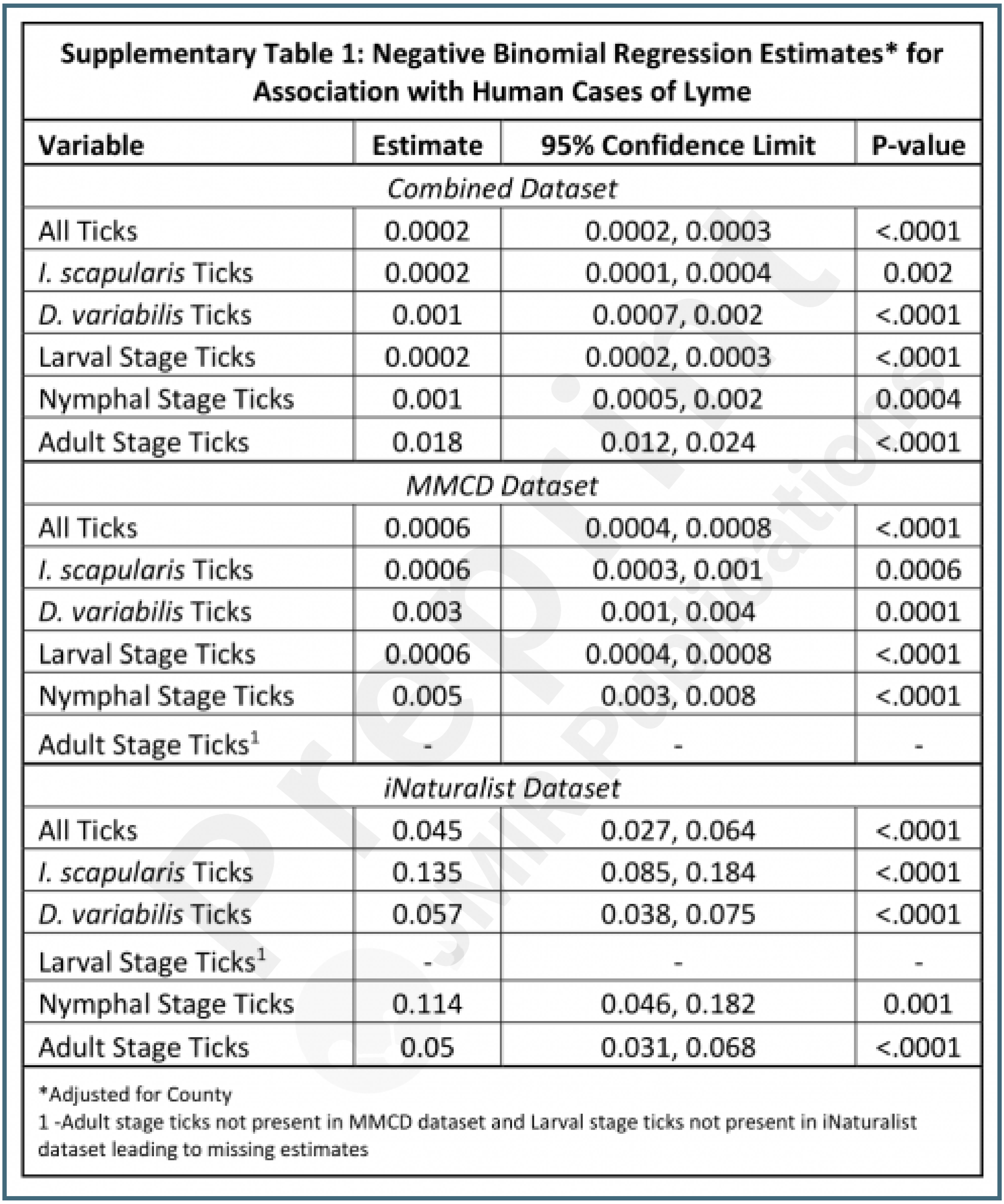
environment at Asian longhorned tick (Haemaphysalis longicornis)-positive sites in Virginia and New Jersey, 2018. Transbound Emerg Dis, 2021. **68**(2): p. 605-614.

16. *Latest annual estimates of Minnesota and it 87 counties' population and households, 2022*. 2022; Available from: https://mn.gov/admin/demography/data-by-topic/population-data/our-estimates.
17. Logan, J.J., et al., *Ixodes scapularis density and Borrelia burgdorferi prevalence along a residential-woodland gradient in a region of emerging Lyme disease risk.* Sci Rep, 2024. **14**(1): p. 13107.
18. Lantos, P.M., et al., *Clinical Practice Guidelines by the Infectious Diseases Society of America (IDSA), American Academy of Neurology (AAN), and American College of Rheumatology (ACR): 2020 Guidelines for the Prevention, Diagnosis and Treatment of Lyme Disease.* Clin Infect Dis, 2021. **72**(1): p. 1-8.

Supplementary Files

Figures

Supplementary table 1 contains negative binomial regression estimates for association with human cases of Lyme disease.



Supplementary Table 1: Negative Binomial Regression Estimates* for Association with Human Cases of Lyme

Variable	Estimate	95% Confidence Limit	P-value
Combined Dataset			
All Ticks	0.0002	0.0002, 0.0003	<.0001
I. scapularis Ticks	0.0002	0.0001, 0.0004	0.002
D. variabilis Ticks	0.001	0.0007, 0.002	<.0001
Larval Stage Ticks	0.0002	0.0002, 0.0003	<.0001
Nymphal Stage Ticks	0.001	0.0005, 0.002	0.0004
Adult Stage Ticks	0.018	0.012, 0.024	<.0001
MMCD Dataset			
All Ticks	0.0006	0.0004, 0.0008	<.0001
I. scapularis Ticks	0.0006	0.0003, 0.001	0.0006
D. variabilis Ticks	0.003	0.001, 0.004	0.0001
Larval Stage Ticks	0.0006	0.0004, 0.0008	<.0001
Nymphal Stage Ticks	0.005	0.003, 0.008	<.0001
Adult Stage Ticks[1]	-	-	-
iNaturalist Dataset			
All Ticks	0.045	0.027, 0.064	<.0001
I. scapularis Ticks	0.135	0.085, 0.184	<.0001
D. variabilis Ticks	0.057	0.038, 0.075	<.0001
Larval Stage Ticks[1]	-	-	-
Nymphal Stage Ticks	0.114	0.046, 0.182	0.001
Adult Stage Ticks	0.05	0.031, 0.068	<.0001

*Adjusted for County

1 -Adult stage ticks not present in MMCD dataset and Larval stage ticks not present in iNaturalist dataset leading to missing estimates