**Title: Community dynamics of ticks from Pennsylvania**

**Damie Pak**

**Steven B. Jacobs**

**Joyce M. Sakamoto**

**Abstract:**

**\begin{abstractbox}**

**\begin{abstract} % abstract**

**\parttitle{Background} %if any, e.g. First part title**

In the last two decades tick-borne diseases have been increasing at the local, national, and global levels. Researchers studying ticks and tick-borne disease need a thorough knowledge of the pathogens, vectors, and epidemiology of disease spread. Three surveillance approaches are commonly used to provide insight into tick-borne disease risk: human disease case surveillance, active tick surveillance, and passive tick surveillance. Passive surveillance alone is a powerful tool for estimating tick distribution, community dynamics, and potential risk of tick bites across a wider geographic area. Long-term passive surveillance can provide up-to-date data on the spatial variability and temporal dynamics of ectoparasite communitiesand shed light into the ecology of rarer tick species. We present a retrospective analysis on compiled data of Pennsylvania ticks (principally consisting of public submissions, but also including the results of active tick surveillance) over the last 117 years.

**\parttitle{Methods} %if any, e.g. First part title**

We compiled data from ticks collected during tick surveillance research, and from citizen-based submissions to the Penn State University Department of Entomology (PSUEnt). Specimens were housed as part of the PSUEnt arthropod collections that eventually became The Frost Entomological Museum. We analyzed all data from 1900-2017 for host associations and spatio-temporal changes

**\parttitle{Results} %if any Second part title**

In total there were 4,491 submission packets consisting of 7,132 tick specimens. Twenty-four different species were identified, with the large proportion of submissions represented by five tick species (most notably *Dermacentor variabilis* and *Ixodes scapularis*). We observed a shift in tick community composition in which the dominant species of tick (*I. cookei*) was replaced by *D. variabilis* in the early 1990s, which was then overtaken by *I. scapularis*. We analyzed quantitative host association data and determined that there are overlaps in host range amongst different, suggesting potential hubs of pathogen transfer between different tick vectors and between different reservoir hosts.

**\parttitle{Conclusions} %**

Text for this section.

We highlight the importance of long-term passive tick surveillance in investigating the ecology of both common and rarer tick species. Information on the geographic distribution, host-association, and seasonality of the tick community can help researchers and health-officials in identifying high-risk areas. Passive-surveillance is an inexpensive and effective method to monitor future trends which is pertinent with growing concerns of climate change and increased human development.

**\end{abstract}**

**Keywords**

**Author Contributions**

*Damie Pak*: Data visualization, analysis, writing

*Joyce Sakamoto*: Data organization, analysis, writing.

*Steve Jacobs*: Database creation, metadata collection, and tick identification

**I . Introduction**

The Centers for Disease Control recently reported an increase in vector-borne diseases in the USA, 80% of which are transmitted by ticks (Rosenberg et al., 2018). Although there are multiple serious tick-borne pathogens, the vast majority of tick-borne disease cases are caused by *Borrelia burgdorferi*, the main causal agent of Lyme Disease in the USA. Pennsylvania has had the highest number of total Lyme disease cases since 2000, with increasing numbers of cases across several counties (Figure 1).

The increase in tick-borne disease is attributed to multiple abiotic and biotic factors . Changes in geographic variability and the temporal dynamics of tick species may influence tick-borne disease outcomes. Several tick species have been expanding into new habitats across the North American continent and posing novel risks to local communities (Sonenshine, 2018). While tick species can be introduced through migrating birds or through human activities, recent surveillance reports suggest that some particularly invasive species are endemic .

**Figure 1**: Annual reported cases of Lyme disease by state from 2006-2017 (Left) and the annual reported cases of Lyme diseases by counties in Pennsylvania from 2006-2017. Public data from the Center of Infectious Disease.

One of the strategies employed by vector biologists to detect introduced vector species is vector surveillance. Arthropod vector surveillance can be a powerful tool for the detection of introduced species (transient or established), emergent arthropod-borne pathogens, and disease risks due to increases or changes in vector population composition. Active tick surveillance approaches such as dragging or flagging, or live animal capture, can be very effective for assessing tick load by habitat (Bouchard et al., 2013; Kollars & Oliver, 2003). However, active surveillance is labor-intensive, costly, and difficult to implement over a wide geographic area. Passive surveillance, in which citizens submit ticks for identification and/or pathogen testing, can be more cost-effective and less labor-intensive, providing data on ectoparasite abundance, host associations, or habitat associations across a wider geographic area (Oliver, Bennett, Beati, & Bartholomay, 2017). Although tick samples may be in less-than-ideal conditions (e.g. missing taxonomically diagnostic parts), or accompanied by incomplete metadata, citizen-submitted tick data can accurately represent tick encounter frequency and potential risk of tick-borne disease exposure (Ripoche et al., 2018).

Passive surveillance data collected over decades may reveal spatio-temporal changes in ectoparasite communities. Data such as spatial distribution and occurrence of both abundant and rare species of ticks can be correlated with land-use (e.g. habitat loss, fragmentation, management), fluctuating environmental conditions, or changes in human or animal behavior (e.g. encroachment may bring reservoir hosts such as groundhogs in closer proximity). Additionally, long-term surveillance data can also reveal shifts in temporal dynamics of tick populations and communities (Sonenshine, 2018). While the seasonality of the major tick species have already been described, year-to-year distribution of tick species may be influenced by inter-annual variability in abiotic (example: local climate) and biotic factors (local reservoir species abundance). To create predictive models that accurately measures risks of tick-borne diseases, passive-surveillance provides a plethora of invaluable ecological data.

We present a retrospective analysis of tick collection data from the early 1900s to June of 2017. Samples were submitted to the Penn State University Department of Entomology and much of the collection was later housed in the Frost Entomological Museum. Specimens submitted prior to 1960 reflect the sporadic nature of citizen-based submissions. An active advertisement campaign was used to enlist the help of citizens in the 1960s and again in the early 1990s. Subsequently, post-funding submission rates decreased in volume, but tick submissions are still received regularly by PSU Entomology for identification. This database provided information about the prevalence of the most commonly encountered tick species by host, relative species abundance over time, occurrence by county, and (for some dates) vegetation associated with tick acquisition. We used this dataset to

1) map the distribution of major tick species in PA at the county-level

2) investigate tick community dynamics and seasonality over time, and

3) utilize accompanying metadata to explore tick-host associations and the effect of vegetation on tick species.

**Methods**

**Study locations**

The state of Pennsylvania is located in the mid-Atlantic region of the United States (latitude: 39° to 42°N and longitude: -80° to -74° W). The climate varies across Pennsylvania depending on the region and altitude, but it typically consists of hot, humid summers and winters with heavy snowfalls in certain areas. The majority of Pennsylvania’s land use is dedicated to agriculture (both croplands and pastures), forestland, and dense urban areas. While there have been significant changes in Pennsylvania’s population from 1960 to 2010, a large proportion of the Pennsylvanian population is heavily clustered around two major urban areas, Philadelphia and Pittsburgh, which are located in the south-east corner and in the south-west of Pennsylvania, respectively (Figure 2). While most specimens were collected within the boundaries of the state of Pennsylvania, a few were declared from people either visiting or returning from visiting other states. There were also some tick specimens identified as species that are not commonly found in Pennsylvania and are presumed to have been imported. These non-PA data were not included in state-wide analyses but were included in the supplements (Table S1).

Figure 2: Total population in Pennsylvania counties in 1960, 1990, 2000, and 2010.   
  
Table S1: Tick submissions from outside Pennsylvania.

**Submissions**

The PSU Frost Entomological Museum houses arthropod samples collected by researchers, teaching collections, and samples submitted by the public for identification. Although some of these samples date as far back as the late 1800s, we only present our analysis of the tick specimens from 1900 to June 2017. Because tick samples were submitted to the Frost Museum over a period of 117 years, they represent multiple collection/submission periods (early 1900-1959, 1960s, 1970-1988, Tick Research Lab [TRL] submissions from 1990-1993, and 1995-present).

Data from the ticks submitted prior to 1989 were identified by museum staff and stored in alcohol at the Frost Museum. Subsequent submissions from private citizens or by county extension officials from 1989 until June of 2017 (“Tick Research Lab” [TRL] submissions) were identified by SBJ and stored in alcohol. All data were combined into a single dataset for our analyses. We compiled these data to identify how species abundance have changed over the years. We defined submissions as a vial or packet containing one or more ticks from a given individual. For analysis of host-association, we used “submissions” versus total tick counts by host. We chose to use this more conservative measure rather than the actual numbers of ticks submitted to avoid a skew in abundance by host. For example, a single submission of 1 tick versus 50 ticks from a host were both treated as a single submission. For distribution of tick species over time, we used total tick counts.

**Associated data**

**Identification**

Ticks were morphologically identified to species and life stage using the following taxonomic keys: Ixodidae East of the Mississippi (Keirans and Litwak, 1989), Argasidae (Cooley and Kohls, 1944), *Ixodes* (Keirans and Clifford, 1978), *Dermacentor* (Yunker et al., 1986), nymphal *Ixodes* (Durden and Keirans, 1996), and nymphs of *Amblyomma* (Keirans and Durden, 1998). Identification of samples to species is crucial since at least 3 Dermacentor species, 3 species of Amblyomma, and 9 different *Ixodes* species have been reported in Pennsylvania. If morphologically important characteristics were missing due to damage to the sample, the next level of taxonomic identification was used (e.g. samples with missing mouthparts that were clearly Prostriata were identified as “*Ixodes* spp.”). In a few cases, samples were not able to be identified beyond “tick” and were designated “Ixodidae” or “Argasidae”. Vouchers, difficult-to-identify, or unusual specimens were submitted to the National Tick Collection, Georgia Southern University for confirmation (by Dr. James Oliver at the time of confirmation).

**Spatial distribution of ticks**

To get an overview of the spatial variability of the tick community from 1900 to 2017, we summed the total tick counts by species across the counties. Rarer tick species with less than 150 individual specimens were grouped by genus. We then plotted the total individuals with a dot-density map with each dot representing a single individual. As the associated geographic data for each submission was at the county-level, the placement of the dots within each county's boundary was randomized.

We focused on the geographic distribution of the five most abundant species in our database, which are of significant public health and veterinary importance: *Amblyomma americanum, Dermacentor variabilis, Ixodes cookei, Ixodes scapularis, and Rhipicephalus sanguineus*. Working on the assumption that counties with higher populations would submit more specimens than less populated counties, we estimated the incidence rate or the total numbers of individual ticks per 100,000 people. This was done by adjusting the total tick count of each species by the county's total populationBecause we did not have the population data for each year, we looked at relevant time periods during the surveillance program: 1960-1970, 1990-2000, 2000-2010, and 2010-2020. For each time periods, we used the 1960, 1990, 2000, and 2010 United Census data (CITE) respectively to calculate the incidence rate (Figure 2)

For the tick species with less than 150 submissions across 1900 to 2017, we aggregated the submissions by genus. We excluded the five major tick species (*Amblyomma americanum, Dermacentor variabilis, Ixodes cookei, Ixodes scapularis, and Rhipicephalus sanguineus*) from this analysis. We then mapped presence or absence of each genus by county.

**Temporal analysis**

To investigate how the annual submissions changed over the course of the passive tick surveillance program, we first aggregated all individual tick specimens by year. Because there were few submissions in the beginning of the program, we grouped the annual submissions into decades starting from 1900-1910 (the years included would be from 1900 to 1909) to 2010-2020.

Since the majority of the annual total tick submissions were comprised of *A. americanum, D. variabilis, I. cookei, I. scapularis,* and *R. sanguineus*, we analyzed the temporal dynamics of these top five taxa. We did not evaluate total counts by year as it varied drastically during active campaigning for citizen submissions or introduction of identification fees. Therefore, we looked at the proportional contribution of each species to the annual summed counts of the five major species. To detect if there have been any monotonic trends (i.e. gradual shifts in abundance), we ran a non-parametric, two-sided Mann-Kendall trend test on the yearly proportion of each of the species from 1900-2017 (Kendall, 1975).

**Seasonality**

For analyzing seasonal patterns of submissions in the passive-surveillance program, we looked at the months of when the tick specimens were received for identification. For some submissions, there were no dates of when the ticks were found by the citizen but typically the length of time between the citizen's discovery and when the specimen was received for identification was short. We first looked at the frequency of the months in which the ticks submissions were received for the collective tick community across 1900-2017 and then by decades. We then investigated the seasonality of the five major tick species individually. For some of the submissions, information of the life-stage (larvae, nymphs, and adults) was included. We then explored if the distribution of the monthly submissions for the life-stages for the major tick species.

Tick vegetation

For a small subset of these data (1989-1990, TRL) the vegetation associated with the tick encounter were also recorded. We defined the vegetation types as brush, ecotone, forest, managed, and pasture. We performed a three-way analysis of variance (ANOVA) test to explore if the total tick counts were dependent on the different types of vegetation. For the ANOVAL model we included the total tick count and our different predictors were the 5 vegetation type, the different host species, and the tick-species.

**Host Associations**

Host information was available for many of the tick specimens (combined by Family, except for dog, cat, human, and groundhog). Host data were classified as either domestic or wildlife. We summarized the host-tick data by summing the total submissions by both the tick species and the host groups. We constructed a circular network map to visualize the relationships between tick species and hosts. All data analyses in this paper was done with R (Version 3.4.1, RStudio version 1.1.383) with the packages MannKendall (Mcleod 2015) for the Mann-Kendall test and the circlize package for the network maps (Gu 2013).

**Results**

*General observations*

From 1900 to 2017, PSU Entomology handled a total of 4,491 submission packets consisting of 7,132 tick specimens. Twenty-four species were identified, although there were many specimens (n=355) that were not identifiable to species due to specimen damage (Table 1). There were five species of ticks that accounted for the majority (91%) of the total count: *Dermacentor variabilis* (n = 3172), *Ixodes scapularis* (n = 1899), *Ixodes cookei* (n = 897), *Rhipicephalus sanguineus* (n = 332), and *Amblyomma americanum* (n=196). These five species are of medical and/or veterinary significance as vectors of tick-borne zoonoses including Lyme Disease (LD), tick-borne relapsing fever (*Borrelia miyamotoi*), babesiosis (*Babesia microti)*, anaplasmosis (*Anaplasma phagocytophilum)*, human ehrlichiosis (*Ehrlichia spp*.), Rocky Mountain Spotted Fever (RMSF), and Powassan Encephalitis Virus. Other tick species that were represented in at least 100 submissions were *Dermacentor albipictus* (n = 107), *Ixodes dentatus* (n = 120), and *Ixodes texanus* (n= 111). Soft tick specimens were also submitted, but at low frequencies. This was probably not representative of the soft tick load in Pennsylvania, but rather reflective of the hosts from which they were obtained (birds or bats).

**Table 1:** The total submissions to the Frost Entomological Museum from 1900 to 1970s, and to the Department of Entomology from 1989-2017. We included known hosts from our database. For submissions in which the ticks were not attached to a host, we recorded the host as “unknown”.

**Spatial Analysis:**

**Figure 3: Dot-density map of all individual tick specimens across Pennsylvania from 1900-2017. Each point represents an individual specimen with its placement randomized within the county.**

From 1900-2017, ttick specimens were submitted from all 67 counties in Pennsylvania (Figure 3). We suspected that higher tick submission numbers would come from areas with higher populations and as expected, tick submissions were heavily clustered around Allegheny and Philadelphia County where Pittsburgh and Philadelphia are located respectively. However, when we adjusted the total tick count by county population, we found higher incidence rates in *less* populated counties. For example, in 1990-2000, the highest incidence rates of *Ixodes scapularis* submissions were from Elk County (870 individuals per 100,000 population). The neighboring Forest and Cameron counties also had high submissions of *I. scapularis* with 116.64 and 589.97 individuals per 100,000 respectively. Other counties with high *I. scapularis* incidence rates included Northumberland (299.56 individuals per 100,000), Snyder (198.98 individuals per 100,000), Union (171.31 individuals per 100,000), and Clearfield (99.21 individuals 100,000) (Figure 3).

**Figure 4: Incidence rates of** the five most abundant tick species across Pennsylvania at different time periods from 1960-2020 . On the left, is the cumulative individuals

*Dermacentor variabilis* distribution was largely localized to southern portions of the state, In 1990-2000, the highest proportion of *D. variabilis* submissions came from Greene County, the most southeastern county of Pennsylvania (865.45 submissions per 100,000). Other southern counties with significantly high rates included Fulton County (350.60 per 100,000) and Franklin County (117.26 per 100,000). However, *D. variabilis* seem to be clustered in parts of northern Pennsylvania with the the second highest incidence rates of *D. variabilis* were from Pike County with incidence rates of 688.955 individuals per 100,000.

*Ixodes cookei* was more evenly distributed throughout Pennsylvania, although similar to *Ixodes scapularis*, it was more highly abundant in the northern counties. In 1990-2000, Forest county had the highest incidence rates of *Ixodes cookei* with 80.87 per 100,000. *Rhiphicephalus sanguineus* and *Amblyomma americanum* had very few submissions and their distribution was mostly scattered across Pennsylvania.

Multiple species within the genera *Ixodes* and *Dermacentor* were widely distributed across Pennsylvania (*I. scapularis, I. cookei, D. andersoni*, and *D. albipictus*) (Figure 5). Other species in the genera *Amblyomma, Argas*, *Carios* (*Ornithodoros*), and *Haemaphysalis* were not as widely distributed, possibly because these species are not commonly encountered or because the specimens were introduced from their native geographic ranges. For example, we only had four submissions of *Argas cooleyi* and *A. persicus.*

**Figure 5 .** Presence or absence map of the tick genuses excluding *Ixodes scapularis, Ixodes cookei, Dermacentor variabilis, Amblyomma americanum, and Rhipicephalus sanguineus.*

**Temporal**

**Figure 6: On the left is the** annual sum of tick specimens (log-transformed) from 1900 to 2017. On the right are the proportional contribution of the major tick species to the total tick counts (1900-2017). The grey shaded area represent area where there was no tick submissions.

Figure 7:

***Species abundance***

Prior to the 1990s, the majority of the tick submissions were made up of *I. cookei* and *R. sanguineus* specimens (Figure 6). The spike in the number of submissions 1990 were largely due to *D. variabilis*, but gradually, *I. scapularis* became the dominant taxon submitted. Results from the Mann-Kendall test supports these observations with an upward trend in the *I. scapularis* counts (tau = 0.288, p = 0.02) and a significant downward trend in *D. variabilis* (tau = -0.408, p = 0.002). The Mann-Kendall also indicate that the proportional contributions of *I. cookei* (tau = -0.607, p < 0.001) and *R. sanguineus* (tau = -0.377, p = 0.005) to the total count have also significantly shifted over a century.

**Seasonality**

**Figure 8:** On the left is the total proportion of tick specimens received at different months of the years from 1900 to 2017. On the right are the proportion of the five major tick species received at different months of the years from 1900-2017.

**Figure 9: The monthly proportion of the tick specimens received at different months of the years aggregated at different decades.**

Overall, we find that the majority of tick specimens were received in the months between April and July with May being the month with the highest proportion of tick submissions (Figure 8). However, there was significant variation in the seasonality of the submissions during the program (Figure 9).

The seasonality data for the five most abundant tick species tracks with previous records of seasonality described by other researcher (refs). Submissions of *D. variabilis, Am. americanum, I. cookei and R. sanguineus* were most abundant during the period between May and July. By species, *D. variabilis and Am. americanum* were most abundant from March to October. *I. cookei and R. sanguineus* samples were submitted throughout the year, but had peak abundance in June. Samples of *I. scapularis* were also submitted year-round, but the peak abundances were bimodally distributed, with a large peak in between May to June, and a second peak between October to November. Our data tracks these peaks, although nymphal *I. scapularis* were also submitted during late spring to early fall.

*Life stage abundance by species*

There were 6,233 specimens in which the lifestage of the specimen were known. 4 % of the submissions were identified as larvae (n=237) , 20% of the submissions were identified to nymphs (n = 1271), and 75\% of the submissions were adults (n = 4725 ). Of the five species, only *I. cookei, I. scapularis, and D. variabilis* had enough submissions to make any conclusive statements about their seasonality.

Figure 10 : **The monthly proportion of *D. variabilis* specimens at the larvae, nymphal, and adult stages across 1960 to 2010**

For *Dermacentor variabilis*, the submissions included 32 larvae, 33 nymphs, and 3059 adults from 1960 to 2017. While we only had larval specimens from 1990-2000, we found that the peak submissions was in September (90%). For the nymphal submissions, there was a more unimodal distribution with the peak centered around June. Finally, the adult submissions of both 1990-2000 and 2000-2010 showed similar patterns with the peak in submissions between May and June.

Figure 11: **The monthly proportion of *D. variabilis* specimens at the larvae, nymphal, and adult stages across 1990 to 2017**

Before 1990, there were only six submissions of adult *Ixodes scapularis*. In 1990-2000, the nymphal and larval submissions show a unimodal pattern with the highest proportion of submissions received in June. For the adult submissions during this decade, there are prominent bimodal peaks in May and October with similar proportion of submissions received in both seasons. There were fewer submissions in 2000-2010 with only 23 larvae, 2 nymphs, and 486 adults. However, the monthly submissions of both the larvae and adults in this decade were consistent to the seasonal patterns found in 1990-2000. Finally, in 2010-2020, the adult submissions (n = 31) show a shift with a number of submissions found earlier in March.

Figure 12: **The monthly proportion of *D. variabilis* specimens at the larvae, nymphal, and adult stages across 1980 to 2017**

The majority of *Ixodes cookei* submissions were identified as nymphs with a total of 521 submissions followed by adult (n = 182) and larvae (n= 88). Submission patterns indicate that Ixodes cookei are found all year-round specifically in the nymph. Across all life-stages, we see that the distributions are unimodal with peaks in early summer between May and June.

**Host data: The results section here is missing**

**Vegetation data:**

There were 677 submissions for which we had data on tick associations with vegetation type(from 1989 to 1990). The large proportion of vegetation data was of ecotone (n = 190) which made up 27 percent of the data. This was followed by forest (n = 178) and managed (n=176). The ANOVA for the tick counts indicate that both the habitation and host types were the significant predictors for the tick counts (F = 4.23, p = 0.002 ; F =5.62, p < 0 .001). In the pairwise Tukey comparisons in the vegetation types, the greatest difference in the vegetation type was between managed and ecotone.

**Discussion:**

**Discussion  
1) Spatial stuff:**

1. **Here we have evidence that passive surveillance is far reaching with all 67 counties being accounted for (Figure 3) and when we adjust for population level, there are less-populated counties that are at higher risks.**
2. **Each of the five species have their own habitat suitability, the strongest indication would be between I. scapularis and D. variabilis**

**2) Temporal stuff:**

**A) Here we show strong temporal changes in the submission numbers that could be due to the availability of public funds- few sentence here.**

**B) Shows that because of the long nature of the passive surveillance, we have seen that there has been changes in the tick community specifically with the species dominance shift**

**C) then segway into how changes in the seasonality of submission could be due to the species dominance shift. Specifically, May was when the highest proportion of Dermacentor variabilis specimens were received which reflected that in the overall submissions, but because of the increase of I. scapularis, we might now see a more bimodal distrubtion in seasonality. Maybe clue into how this distribution could change even further with warming springs and winters**

**3) Maybe talk about host associaton and vegetation data which are rare. Maybe not enough data to make any grand-standing conclusions, but is kinda a guide for future research. We can see that we have ignored some species like Ixodes cookei but it may be an important key connecting domestic and wildlife hosts together. The vegetation type is also a key thing with increased development of land.**

**4) Then a short paragraph talking about the importance of citizen science/public engagement and the importance of keeping the passive surveillance going.**

*Importance of tick species identification*

In 2000, Merten and Durden described 84 total tick species naturally occurring in the United States, 40 species of which will bite humans (11 species of soft ticks and 29 species of hard ticks). Of those, 7 species are listed as important disease vectors of humans on the CDC website: 1) American dog tick (*Dermacentor variabilis*), 2) Blacklegged tick (*Ixodes scapularis*), 3) Brown dog tick (*Rhipicephalus sanguineus*), 4) Gulf Coast tick (*Amblyomma maculatum*), 5) Lone star tick (*Amblyomma americanum*), 6) Rocky Mountain wood tick (*Dermacentor andersoni*), and 7) Western blacklegged tick (*Ixodes pacificus*). The tick species of most epidemiological importance in Pennsylvania are *Ixodes scapularis*, the vector of *Borrelia burgdorferi, Babesia microti, Anaplasma phagocytophilum,* andPowassan Encephalitis Virus. Distinguishing this tick genus from other genera of ticks is fairly simple, but *species* level identification requires close morphological scrutiny, since there are 7 other native Ixodes species that appear superficially indistinguishable to the untrained eye. Many of these other species are not competent vectors of *B. burgdorferi*, but may serve as vectors of other pathogens/parasites. It is therefore important to correctly identify tick species, not only for determination of disease risk, but also because the treatments may differ significantly.

*Hidden gems in retrospective data*

Our data encompasses a large timeframe and includes several collection periods. While the promise of such a rich dataset is enticing, it was a laborious task to clean and standardize the data into a format that could be analyzed. Additionally, without context, it was difficult to understand how nonnative specimens came into the Commonweath of Pennsylvania. Much of the literature on the distribution of ticks in Pennsylvania was found in lists published either in the 1940s or after the 1990s.

Fortunately, we were able to locate a USDA report on ticks and tick-borne disease by Dr. Robert Snetsinger (Snetsinger 1968). This report contained explanations for many of the unusual specimens, but it also provided a glimpse into how much change has occurred since 1968 in the PA tick community. Although the report did not have actual counts (only percentages) with associated metadata (and in fact contained references to data from previous papers or unpublished data), we were able to cross-reference the existing tick specimens from this time period with associated data contained in handwritten notes found at the Frost.

We were able to detect population shifts in retrospective analyses, but the Snetsinger report provided insight into what was known at the time about many of the tick species found and what pathogens they could spread. For instance, it was known, but not published, that the deadly, but rare Powassan Encephalitis Virus (PEV) was present in groundhogs in southern PA. The most important tick-borne disease at that time was spotted fever (chiefly *Rickettsia rickettsii*, although there is some evidence that other spotted fever rickettsias may also be in Pennsylvania). *R. rickettsii* is transmitted principally by *D. variabilis*, but since then has been shown to be transmitted by *R. sanguineus* (ref).

Finding the Snetsinger report was a boon for us, particularly because it provided the answers to why so many exotic tick species were present in the Commonwealth. For example, the specimen of *Amblyomma cajennense* (the Cayenne tick generally limited to neotropic regions) was collected from a capybara in 1913 from the Philadelphia Zoological Garden (Snetsinger 1968). A century later, Brazilian researchers documented spotted fever-infected *A. cajennense* on capybaras (Krawczak et al., 2014). Perhaps if this specimen has been maintained well enough, we might be able to extract DNA and assay for spotted fever rickettsial species.

Other unusual specimens included a single European species of tick, *Ixodes ricinus* (the castor bean tick), which is sister taxon to *Ixodes scapularis*, was found on lizards in the Pittsburgh Zoo (Snetsinger 1968). Another reptile-associated tick *A. dissimile* was known to be imported through the pet trade (snakes) and on research animals. In one humorous, but almost hidden anecdote, one of the tick specimens came from a snake that was used as part of the costume by a “night club ballerina” (Snetsinger 1968). Data on three other rare species were found on birds (*I. brunneus, Haemaphysalis chordeilis,* and *A. longirostre)*(Snetsinger 1968)*.*

There was one specimen of *Amblyomma maculatum*, found in the ear of a child from Pittsburgh that had never left the state (Snetsinger 1968). In the 1960s, this species was only known from the neotropics and nearctic regions along the coasts of southern Atlantic states, so the presence of this specimen was “a mystery” (Snetsinger 1968). In the last several decades, the range of this species has expanded westward into the Mid-west and northward into the Mid-Atlantic on migrating birds (Sonenshine 2018). Their establishment northward has been attributed to climate change, but because they require high humidity and higher temperatures, what has driven their movement inland and westward is not yet clear. It is possible that successful breeding populations are more tolerant of cooler and/or drier conditions, or alternatively, because they are localizing to warmer and/or more humid areas along riparian ecotones (Sonenshine 2018).

*Argas persicus* is an Old World species associated with poultry, and has rarely been collected in the United States, with reported sightings confirmed by Glen Kohls from California, Georgia, Maryland, and Pennsylvania (Kohls et al., 1970; Muñoz-Leal et al., 2018; Snetsinger, 1968).The two confirmed instances of *A. persicus* from Pennsylvania chickens came from York and Adams county in (FIND YEAR) (Kohls et al., 1970).

*Argas cooleyi* is mostly associated with cliff swallows in southwestern states of the USA, so humans do not normally encounter them unless they disturb the nests or if the nests are in close proximity to human dwellings (Beatty et al., 2017). It is therefore unlikely that this specimen is commonly encountered in Pennsylvania unless imported by humans or through bird movement.

*Comparisons of tick communities from 1960 to present*

In the 1960s the tick communities were dominated primarily by three species: *Ixodes cookei*, *Dermacentor variabilis*, and *Rhipicephalus sanguineus*. Even at this *time Ixodes cookei* was described on as many as 10 different hosts including humans and dogs. The second most abundant species*, D. variabilis*, was widely distributed and eventually overtook *I. cookei* in abundance in the 1990s. This was subsequently overtaken by *I. scapularis* abundance.

The third most abundant species of note in the Snetsinger report was the Brown Dog tick. *Rhipicephalus sanguineus* originated in Africa, but has since become a cosmopolitan urban pest species found worldwide in association with humans and their canine companions (Brites-Neto et al., 2015). In 1968, Snetsinger suggested that *R. sanguineus* had established breeding populations in Pennsylvania, but according to our records, we have had very few submitted during the TRL years and none since 2002. Since both *R. sanguineus* and *D. variabilis* are competent vectors of Rocky Mountain Spotted Fever (RMSF), the recognition of either of these species is important for diagnosis and treatment after a bite (Promed Archive Number: 20140908.2757657).

Lastly, the presence of *Amblyomma maculatum* and *A. americanum* were very rare in the 1960s, but have become more common since. *A. maculatum* is a known vector of *[Rickettsia parkeri](https://en.wikipedia.org/wiki/Rickettsia_parkeri" \o ")* (ref). *A. americanum* is either a known vector or competent vector of ehrlichiosis, tularemia, multiple *Borrelia species* (Including *B. lonestari, B. andersonii, and B. americanum*), and competent vector of heartland virus)refs. *A. americanum* has also been implicated in inducing meat allergies in certain people (refs). The role of both species as vectors suggest that the continued monitoring for this species is in order.

In terms of seasonality

*Quantifying host association*

Much of the data on tick host associations has been based on studies compiling presence or absence of tick species in museum collections around the country, or identified off of vertebrate museum specimens (Persing et al., 1990). However, in many cases, these associations were published before the 1950s and even in the Snetsinger report, there were no quantitative data available. We were fortunate enough to have access to specimens and catalogued these data into a quantifiable dataset. We were able to confirm host associations described previously, but we also provide counts.

One particular tick-host association of interest is that of *Ixodes cookei*. In the older tick literature, *I. cookei* was considered a cosmopolitan species, parasitizing many medium-sized mammals. When we look the tabular data of host associations (Figure #), we can see that this species, *like I. scapularis, D. variabilis,* and *A. americanum*, does indeed have a wide host range. In some current online tick resources, this species, current online tick resources often suggest that this species is a groundhog or woodchuck tick, not an important vector of *B. burgdorferi*, and therefore not vector of public health interest, apart from rare transmission of PEV).

Thus, *I. cookei* is a largely ignored species for which little is known about its microbiota (including pathogens, commensals, or symbionts). Yet, because of its broad host range and co-parasitization with other tick species, *I. cookei* may be a potential bridge vector that can transmit pathogens from one reservoir host to another. Its natural history and co-parasitization of vertebrates hosts with other species might warrant further scruntiny.

Other host associations should also be more closely examined. There was a recent study suggesting that pet ownership increased the likelihood of tick bites (Jones et al., 2018), but when reviewing the literature, we find that this is not a new observations, nor a surprising one. In 1968, at least six of the 20 tick species described were likely to come into contact with humans through their canine companions (Snetsinger, 1968).

*Passive surveillance and the value of citizen contribution*

The ideal surveillance program for vector-borne disease epidemiology would have both active and passive collection strategies. Combined with reported human cases of tick-borne disease, both passive and active collections provide complementary data for accurate assessment of tick-borne disease risk. The metadata associated with both active and passive tick surveillance (assuming it has been curated and well-managed) can provide insight into tick-host association, vegetation, seasonality, and shifts in population structure that can be used for modeling disease risk. Archival tick samples (or their DNA) can be useful for retrospective mining for research on the population genetics of ticks to detect gene flow, host shifts, or on their microbial inhabitants.

Active collection methods often provide a better estimate of the tick population sizes, a major variable used to assess tick-borne disease risk. Springer et al. described the distribution of lonestar ticks (*A. americanum*) across the United States, but only identified it in 17 of 77 counties of Oklahoma (Springer et al., 2014). Barrett suggested that this was an underestimation due to incomplete passive surveillance (Barrett et al., 2015). Barrett et al (2015) used a combination of retrospective literature review, data compilation of specimens from archival collections, *and* active collection (dry ice, dragging, and flagging) in counties presumed to be free of *Amblyomma americanum* and found that 68 of 77 counties of Oklahoma were colonized. One of the limitations of active collection methods is that while overall population sizes may be high, sampling may not reflect population abundance, particularly if the distributions are spotty. Tick populations are not static and may be highly mobile, depending on the host upon which they alight. Nevertheless, active surveillance can fill in the gaps in tick distribution throughout each state.

Passive surveillance, on the other hand, is sometimes criticized for under-representation of certain taxa or bias toward certain host associations. Nevertheless, citizen-submitted tick collections can provide valuable baseline data on prevalence and likelihood of tick encounters (Barrett et al., 2015; Cortinas and Spomer, 2014; Koffi et al., 2012; Nelder et al., 2014; Ripoche et al., 2018; Xu et al., 2016). In fact, there is evidence that passive tick surveillance data is more strongly correlated with reported human cases of tick-borne diseases than active surveillance (Ripoche et al., 2018). A community engagement program that actively recruits ticks submitted by citizens should be coupled with support for a rigorously curated database of tick submission.

With the caveat that our dataset contains gaps in years of submissions, we were able to detect seasonality, shifts in tick community composition, and host associations that have not been well-documented in a quantitative manner. We found that during the 1990s, *Dermacentor variabilis* (the dominant species submitted in 1989-1990) was quickly overtaken as the dominant species by *Ixodes scapularis*. Although we cannot directly infer a causal negative relationship between these two species from our data, this observation is supported by other studies of tick species diversity. Ohio’s passive surveillance program (started in 1978) did not detect *Ixodes scapularis* (=*I. dammini*) until 1989. At that time, the dominant species were *D. variabilis* (~97% of submissions) and *Ixodes cookei* (1.2%) (Pretzman et al 1990). From 1989 until 2008, *Ixodes scapularis* accounted for less than 1% of the total submissions, but from 2009, the abundance began to increase until in 2012, they accounted for 24.8% of ticks submitted to the Ohio Department of Health (Wang et al 2014).

When we combine citizen-based passive surveillance data with museum collection data, we find that hybrid datasets are a powerful tool for mining past ecological and epidemiological events. Many states maintain county records on passive tick submissions to veterinary or medical health officials, but there may be other cryptic collections (and associated data) housed in museums, universities, government institutions, or with private individuals. We have hopes to combine our data with those in the literature and with currently unexplored collections into a massive meta-analysis. These collections in total will provide the basis for exploring hypotheses such as: 1) are shifts in tick populations correlated with increasing human encroachment on natural habitats, 2) what are some phenological reasons for the increase in *Ixodes scapularis* abundance; or 3) if displacement of a dominant tick community species occurs, what are the implications for tick-borne disease risk? (Johnson et al 2004; Rand et al 2007; )

*Impacts of pro bono tick identification versus charged on submission rates*

The submission rates of ticks from the public were both influenced by thelevel of publicity advertising the tick identification service and shift from free tick identification services to a pay-by-submission model. In the case of the 1960s Snetsinger campaign, submission rates were high largely due to a multimedia advertisement campaign that targeted principally housewives, although other citizens also submitted specimens during that period. Similarly, during the first year of the TRL campaign, there were thousands of submissions, but the submission rates declined over the next two years. Subsequent lack of funding to support f tick species identification meant that to continue to provide the service, the identifications had to be supported by charging by submission. This meant that submissions were mostly limited to those willing to pay 25 dollars per specimen, or to medical professionals who removed the ticks from patients. In addition to the cost of tick identification, there was no available funding for either standardized pathogen screening or database management, which greatly hampered efforts to provide a more robust dataset.

This is in contrast to states such as Maine, where the state Department of Health was directly involved in collection of tick data and screening for tick-borne pathogens (Rand et al., 2007). In just under 2 decades, the Maine DOH program (after engaging in a sustained outreach campaign to encourage public submissions) obtained 24,519 ticks were submitted for identification free of charge. We can only speculate, but predict that the submission numbers to the TRL would have been much higher had the costs of maintaining an active tick identification program been defrayed by an external source rather than setting a fee per tick.

Despite the rise in tick-borne diseases, the movement on the policy front has been slow. Unlike Maine, PA has been reluctant to fund a state-wide tick surveillance program. Perhaps this is because its constituents prefer action against more dramatic mosquito borne diseases such as West Nile or Zika. Pennsylvania has had 539 cumulative cases of West Nile (1999-2016) and 183 cumulative cases of imported Zika (2015-2018). In contrast, since 2006 there have been 48,626 cumulative confirmed cases of Lyme Disease. There is public interest in funding to combat Lyme Disease and tick-borne diseases in PA. In 2016 two of the authors (SBJ and JMS) were contacted by the PA Lyme Disease Task Force to generate a line item budget for a tick surveillance. As of 2018, however, there is still no funded tick surveillance program. This may change, particularly with the identification of a new invasive (and potentially parthenogenic) exotic Longhorned tick (*Haemaphysalis longicornis*) in nearby states. It may only be a matter of time before this vector is found in PA (Rainey et al 2018; Promed 20180603.5836650). When that happens, the absence of a coordinated statewide tick surveillance program may be recognized and identified as a public health priority.

*Conclusion*

Collectively, these data represent the value of museum ectoparasite collections, passive surveillance datasets, and the power of effective data visualization to address questions about ectoparasite community dynamics. In order to prevent loss of this particular cryptic collection of ectoparasite data, we are contributing to a multi-institutional effort to build a better understanding of tick-host associations, interactions and preserve these data for future studies (via Terrestrial Parasite Tracker, <http://parasitetracker.org>). Cryptic datasets such as these exist throughout the United States in museums, public health repositories, and private collections. Arthropod specimens (and/or their metadata) may exist in cataloged and annotated form, or perhaps have been left unsorted and undocumented. Destruction of such entomological collections reflects a lost piece of historical documentation that may lend insight into past epidemiological outbreaks or explain changes in ectoparasite communities. Our data is a contribution to the overall understanding of the historical dynamics of tick communities in Pennsylvania and provides insight into ecological factors that may have influenced spatial and temporal shifts in tick community composition.

*Availability of data and materials section*

The dataset(s) supporting the conclusions of this article is(are) available in the passive\_tick\_surveillance\_2018repository,https://github.com/pakdamie/passive\_surveillance\_tick\_2018/tree/master/MAIN\_DAT