WHAT IS THE BEST PREDICTOR OF ANNUAL LYME DISEASE INCIDENCE: WEATHER, MICE, OR ACORNS?

ERIC M. SCHAUBER, 1,4 RICHARD S. OSTFELD, 2 AND ANDREW S. EVANS, JR.3

¹Cooperative Wildlife Research Laboratory and Department of Zoology, Southern Illinois University, Carbondale, Illinois 62901 USA

²Institute of Ecosystem Studies, Millbrook, New York 12545 USA

³Dutchess County Department of Health, Poughkeepsie, New York 12601 USA

Abstract. Predicting fluctuations in annual risk of Lyme disease would be useful in focusing public health efforts. However, several competing hypotheses have been proposed that point to weather variables, acorn production, or mouse abundance as important predictors of Lyme disease risk. We compared the ability of acorn production, mouse density, and four relevant weather variables to predict annual Lyme disease incidence (detrended) between 1992 and 2002 for Dutchess County, New York, and seven states in the northeastern United States. Acorn production and mouse abundance measured in Dutchess County were the strongest predictors ($r \ge 0.78$) of Dutchess County Lyme disease incidence, but the increase in mouse abundance from 1991 to 1992 was contrary to a decrease in Lyme disease in the following years. The Palmer Hydrologic Drought Index (PHDI) was a significant positive predictor of Lyme disease incidence two years later for three states (0.58 $\leq r \leq$ 0.88), but summer precipitation was generally negatively correlated with Lyme disease incidence the next year $(-0.79 \le r \le 0.02)$. Mean temperatures for the prior winter or summer showed weak or inconsistent correlations with Lyme disease incidence. In four states, no variable was a statistically significant predictor of Lyme disease incidence. Synchrony in Lyme disease incidence between pairs of states was not significantly concordant with synchrony in any weather variable that we examined $(0.02 \le r \le 0.21)$. We found that acorns and mice were strong predictors of Dutchess County Lyme disease incidence, but their predictive power appeared to be weaker spatially. Moreover, evidence was weak for causal relationships between Lyme disease incidence and the weather variables that we tested. Reliable prediction of Lyme disease incidence may require the identification of new predictors or combinations of biotic and abiotic predictors and may be limited to local

Key words: Borrelia burgdorferi; Ixodes scapularis; Lyme disease; mammals; masting; oak; Peromyscus leucopus; prediction; ticks; vector-borne disease; weather.

Introduction

Lyme disease is the most prevalent vector-borne disease in North America, Europe, and parts of Asia, so there is intense public interest in understanding how the risk of exposure varies in space and time. Annual incidence of reported Lyme disease cases in the United States has generally increased over time since the disease was first described, probably due to both expanding populations of tick vectors and increased rates of diagnosis and reporting (Orloski et al. 2000). Lyme disease incidence has also fluctuated from year to year relative to the overall trend, and these fluctuations tend to be consistent across localities within a region (Stafford et al. 1998). Such spatiotemporal consistency suggests that year-to-year fluctuations in Lyme disease incidence in different sites may share an underlying cause

Manuscript received 12 November 2003; revised 4 June 2004; accepted 30 July 2004; final version received 26 August 2004. Corresponding Editor: J. A. Logan.

⁴ E-mail: schauber@siu.edu

or suite of causes. Thus, identifying the causal factors may enable prediction of relative risk.

Lyme disease risk is dependent on the density of blacklegged ticks (Ixodes scapularis) that are infected with the Lyme disease bacterium, Borrelia burgdorferi (Falco and Fish 1989, Fish 1993, Mather et al. 1996, Stafford et al. 1998, Falco et al. 1999). Therefore, fluctuations in Lyme disease risk have been hypothesized to be the result of either fluctuating weather conditions affecting tick survival (Jones and Kitron 2000, Subak 2003) or fluctuating acorn production affecting the behavior and abundance of mammals that act as tick hosts and reservoirs of B. burgdorferi (Ostfeld et al. 1996a, Ostfeld 1997, Jones et al. 1998). Blacklegged ticks are susceptible to desiccation when they quest for hosts (Yoder and Spielman 1992, Stafford 1994, Lindsay et al. 1998, 1999, Vail and Smith 2002), so temperature and precipitation are clear candidates to explain variations in tick abundance and, consequently, Lyme disease risk. Jones and Kitron (2000) observed that the abundance of larval (first-year) blacklegged ticks over eight years in northwest Illinois was positively correlated with summer rainfall and negatively correlated with summer temperatures in the previous year, whereas abundance of nymphs (second year) was positively correlated with prior-year temperatures. However, the abundance of infected ticks may also be mediated by the behavior and abundance of wild hosts.

576

The movements and abundance of mammals that serve as reservoirs of B. burgdorferi and hosts for I. scapularis are profoundly affected by the abundance of tree seeds, especially acorns (Ostfeld 1997, Ostfeld et al. 2001). Acorn production fluctuates greatly from year to year, with synchrony (defined as positive crosscorrelation between concurrent time series from separate sites; Bjørnstad et al. 1999, Koenig 1999) over spatial scales on the order of tens to hundreds of kilometers (Koenig and Knops 2000, Liebhold et al. 2004). A large acorn crop has two direct ecological effects that ultimately lead to indirect effects on Lyme disease risk. Almost immediately, abundant acorns attract white-tailed deer (Odocoileus virginianus), the primary host for adult I. scapularis, into oak forest stands in autumn (McShea and Schwede 1993). The result is an outbreak of larval ticks in oak stands in the following summer (Ostfeld 1997, Jones et al. 1998). Abundant acorns also cause increased abundance of white-footed mice (Peromyscus leucopus) and eastern chipmunks (Tamias striatus) in the following summer (Elkinton et al. 1996, Wolff 1996, Jones et al. 1998); both species are important hosts for larval I. scapularis and are highly competent reservoirs of B. burgdorferi (Donahue et al. 1987, Mather et al. 1989, Slajchert et al. 1997, Schmidt et al. 1999, LoGiudice et al. 2003). Thus, large acorn crops appear to cause concentrations of questing larval ticks and of competent reservoir hosts to coincide in time and space in the following summer, culminating in high density of infected nymphs in the second summer (Ostfeld et al. 2001).

Weather may also affect Lyme disease risk indirectly via ecological effects on hosts. Subak (2003) found that the incidence of Lyme disease in the northeast United States between 1993 and 2000 was negatively correlated with drought conditions two years before and positively correlated with temperatures in the previous winter. These results were interpreted as a combination of direct effects of drought on tick survival and indirect effects of severe winters on the abundance of white-footed mice. However, there is only weak evidence that winter temperatures affect mouse abundance (Lewellen and Vessey 1998, Subak 2003), relative to the high-amplitude fluctuations driven by acorn production.

Because purported causes like weather and acorn production cannot be experimentally manipulated at scales large enough to substantially affect Lyme disease incidence in humans, causality must be cautiously inferred from correlative studies at large scales buttressed by smaller scale experiments that test proposed mechanisms. Statistical analyses are confounded by the fact that both Lyme disease incidence and hypothesized

causes are autocorrelated in space and time. Thus, separate years and sites do not provide completely independent data, and spurious correlations between purported causes and effects can occur by chance much more frequently than otherwise would be expected (Abraham and Ledolter 1983, Koenig 1999). Spurious correlations are especially problematic because reliable Lyme disease data are only available for a small number of years (since 1993) and because Lyme disease incidence has had a consistent upward trend during that time period.

The search for general explanations may be doomed if different factors govern Lyme disease risk in different places. However, our approach was to ascertain initially whether the few variables that have been supported as universal or nearly universal drivers are able to explain patterns of Lyme disease incidence. If they fail, perhaps we must resign ourselves to the hypothesis that all is spatially idiosyncratic. The search for causes of Lyme disease fluctuations has also been hampered by a tendency of researchers to test the plausibility of a single hypothesis, rather than comparing the explanatory power of two or more alternative hypotheses. Only Ostfeld et al. (2001) explicitly confronted competing hypotheses against a common data set, finding that acorn production was a much better predictor of the abundance of infected nymphal ticks in Duchess County, New York, than were prior precipitation and growing degree-days.

Our objective was to use model selection procedures to determine whether Lyme disease incidence was best predicted by weather variables identified in previous studies, acorn production, or mouse abundance. It was not our objective to seek out and identify new hypothesized predictors, so we restricted our analyses to weather variables that have received empirical support in prior studies. Our analysis included the seven northeastern U.S. states examined by Subak (2003). We also focused on Dutchess County, New York, an area where Lyme disease is hyperendemic (7% of national cases in 2002) and where time series of acorn production, mouse abundance, and weather data were available. We further analyzed weather hypotheses by testing for spatiotemporal concordance of weather variables and Lyme disease incidence across states in the northeastern United States.

METHODS

Lyme disease data

Our response variable was the annual Lyme disease incidence (LDI, i.e., number of case reports) during the years 1990–2002, measured for Dutchess County (New York) and the seven northeastern states responsible for the vast majority of Lyme disease cases in the United States: Connecticut, Maryland, Massachusetts, New Jersey, New York, Pennsylvania, and Rhode Island. The use of state-level LDI data could obscure the im-

portance of factors whose importance varies at smaller scales. In most states, however, Lyme disease cases are concentrated in particular hyperendemic areas (e.g., southeastern New York, coastal areas of Massachusetts and Maryland), so factors affecting disease risk in those areas drive temporal patterns in state-level incidence. For example, acorn production is unlikely to explain fluctuations of LDI in northern New York State, where oaks are rare, but this region makes a very small contribution to the total number of cases in the state (New York State Department of Health 2003).

For Dutchess County, we used case surveillance data collected by the Dutchess County Department of Health, arranged by year of diagnosis. These data stemmed from passive case reporting as well as laboratory surveillance (Chow et al. 2002). A backlog of case reports began building in 2000 due to the West Nile Virus outbreak, and some reports for 2000, 2001, and 2002 remained pending confirmation as of March 2003. Therefore, we estimated the final number of confirmed cases for each year by multiplying the number of pending cases for that year by the proportion of reviewed cases that were confirmed to meet the case definition over the years 1998-2003 (0.452 \pm 0.004 mean \pm 1 sE). For states, we analyzed case reports submitted to the Centers for Disease Control and Prevention (CDC) by state health departments, arranged by year of report (Centers for Disease Control and Prevention 2002, Centers for Disease Control and Prevention 2003; available online).⁵ Some discrepancies and problems in reporting exist. For example, some cases were not reported to CDC in the year of diagnosis due to processing delays. For Dutchess County we were able to attribute all case reports to the year of diagnosis, resulting in some differences from data reported to CDC. Also, these surveillance data underestimate true LDI (Meek et al. 1996), and the CDC case definition was made more stringent in 1997 (Centers for Disease Control and Prevention 1997), which may have reduced the reporting rate relative to earlier years. Finally, 2002 incidence data are provisional, and are likely to change slightly as further reports are submitted and reviewed.

Acorn data

Production of acorns was measured annually from 1992 to 2002 by the abundance of acorns falling per square meter of ground under oak trees in two 2.25-ha forest plots at the Institute of Ecosystem Studies (IES), Dutchess County, New York. In each plot, 20 0.5-m² circular baskets were placed underneath mature oaks, and intact mature acorns in these baskets were counted every month during autumn. Acorn production was measured by dividing the total acorn count from all baskets by the total basket area. Because Lyme disease risk is expected to be linked to acorn production two years before, our acorn data (Fig. 1A) could be com-

pared with Lyme disease incidence in the years 1994–2002.

Mouse data

The hypothesized effect of acorn production on Lyme disease risk is, in part, due to effects on abundance of white-footed mice, a primary reservoir host. Populations of white-footed mice at IES have been monitored by capture-mark-recapture methods since 1991 on two 2.25-ha trapping grids in oak-dominated forest habitats (Fig. 1A). Capture and handling of small mammals conformed to Institutional Animal Care and Use Protocols issued annually. Each year, small mammals were live-trapped for 2-3 consecutive days every 3-6 weeks. Trapping generally began in April or May and ended in November. Each trapping grid consisted of an 11 × 11 array of trap stations, with two Sherman live traps (H. B. Sherman Traps, Tallahassee, Florida, USA) at each station. Traps were baited with oats or sunflower seeds, and cotton batting was provided as insulation during cool weather. Traps were opened in the evening and checked and closed the following morning. Each captured animal was marked with a uniquely numbered ear tag and released at the site of capture. Because each trap session was only two nights, we estimated mouse abundance in each trap session by inputting data from all trap sessions in a year into the Jolly-Seber open population model in program PO-PAN5 (Arnason and Schwartz 1999). Larval ticks in the northeastern United States feed on mice primarily in late summer (Fish 1993), and larval tick densities at IES have generally peaked between mid-August and early September each year (Ostfeld et al. 1995, 1996a, b). Therefore, we estimated mouse abundance on 15 August of each year by linear interpolation between mouse abundance estimates for trap sessions immediately before and after 15 August.

Weather data

Because an enormous number of weather variables could conceivably affect Lyme disease risk, there is great danger of uncovering statistically significant, but spurious, relationships when testing a large number of explanatory variables without clear a priori justification (Anderson et al. 2000). Therefore, we restricted our weather variables to those that have been empirically linked to changes in abundance of immature blacklegged ticks or Lyme disease incidence in prior studies. Based on hypotheses presented in prior studies (Jones and Kitron 2000, Subak 2003), we evaluated the predictive power of four explanatory variables calculated from lagged (1-2 years before incidence data) weather measurements: the mean summer (June-August) Palmer Hydrologic Drought Index lag 2 (PHDI2), mean winter (December-February) temperature lag 1 (MWT1), mean total summer precipitation lag 1 (TSP1), and mean summer temperature lag 1 (MST1). These data were downloaded from the National Climatic Data Cen-

⁵ (http://www.cdc.gov/ncidod/ovbid/lyme/epi.htm)



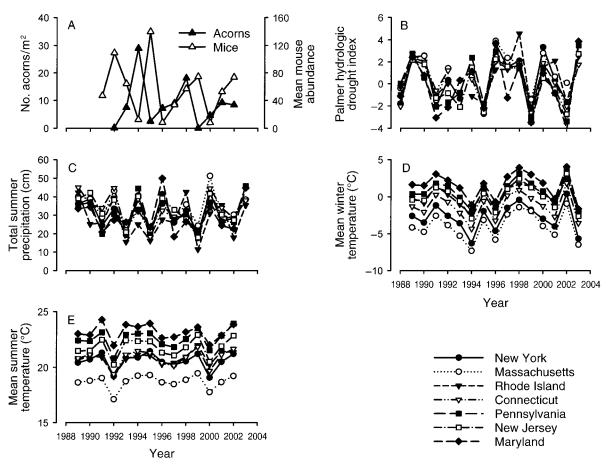


Fig. 1. Time series of variables used to predict Lyme disease incidence: (A) acorn production and mouse abundance measured in Dutchess County New York; (B–E) weather variables measured in each of seven states in the northeastern United States.

ter web site.6 Although Jones and Kitron (2000) found a possible relationship between nymphal abundance and summer growing degree-days in the prior year, we chose to use mean summer temperature instead because growing degree-day data were much more difficult to interpolate across climatic regions. Subak (2003) found weak support for the hypothesis that PHDI is consistently linked with LDI in the same year, so we did not include unlagged PHDI among our potential predictors. Jones and Kitron (2000) claimed in their abstract that cumulative degree-days were correlated with immature blacklegged tick abundance in the same year, but this relationship was not mentioned in their results, so we did not include unlagged temperatures. Based on Subak (2003) and Jones and Kitron (2000), we expected a priori that Lyme disease risk would be positively correlated with each of these weather variables except TSP1. High TSP1 could enhance Lyme disease risk by reducing desiccation of larval ticks, but it could also possibly foster fungi and other natural enemies of ticks (Samish and Rehacek 1999). Unlike acorn and mouse

⁶ (http://www.ncdc.noaa.gov/oa/climate/climatedata.html)

data, which were recorded at a single Dutchess County site, weather data were available for each state individually (Fig. 1B–E). Following Subak (2003), we calculated each variable for the climatic region of each state where the bulk of Lyme cases occurred: Maryland Region 6 (eastern), Massachusetts Region 1 (eastern), New Jersey Region 1 (northwestern), New York Region 5 (Hudson Valley), Pennsylvania Region 3 (southeastern), and Rhode Island Region 1 (entire state). We averaged data from the three climatic regions in Connecticut because recent cases have been distributed throughout the state.

Statistical analysis

For each state and Dutchess County, we detrended annual LDI data from years 1990 to 2002 by applying a natural-logarithm transformation, fitting a linear or quadratic ($y = a + b*[year - 1989] + c*[year - 1989]^2$) trend model, and retaining the residuals. For each data set, the choice of linear or quadratic trend was based on the Akaike information criterion adjusted for small samples (AIC_c; Burnham and Anderson 1998). We then calculated the temporal cross-correla-

tion between detrended LDI and each explanatory variable singly (Bjørnstad et al. 1999). An alternative analysis would have been to incorporate trend components and hypothesized causal variables simultaneously in multiple regression models. However, this would require ignoring trend information contained in data from years in which all explanatory variables were not measured. We considered accounting for the biennial nature of the *I. scapularis* life cycle in the detrending process by adding an even-year indicator variable. We chose not to because TSP also exhibited a tendency for two-year oscillations (Fig. 1C), so removing differences between odd- and even-year LDI would probably obscure any relationship with TSP.

We compared the strength of various predictors by measuring their cross-correlations with LDI. Because acorn data could only be compared with LDI from 1994 to 2002, we compared the cross-correlations with LDI and each explanatory variable over those years. Mouse abundance could be compared with LDI over the years 1992–2002, so we compared the cross-correlations of LDI during that period with all explanatory variables except acorns. We assessed statistical significance of cross-correlations after adjusting degrees of freedom for serial autocorrelation (Sciremammano 1979). Because previous studies allow a priori expectation of a positive relationship, we used one-tailed *P* values for all explanatory variables except TSP1.

After comparing the predictive power of each explanatory variable individually, we then tested whether combinations of variables would improve prediction. Because there is no a priori basis for choosing particular variable combinations, we chose the most parsimonious (lowest AIC_c) main-effects model out of all possible subsets, ranging from the null (intercept only) model to a model with all six explanatory variables. The results of this exploratory analysis must be interpreted cautiously, because the potential for spurious results is high (Anderson et al. 2000). However, predicting fluctuations in LDI could require more than one causal explanatory variable.

Because weather data were available for each state, we were able to test for spatiotemporal concordance between detrended LDI and each weather variable. This procedure, which has been used to test purported causes of mast fluctuations (Schauber et al. 2002), goes beyond the temporal correlation between purported cause and effect at a given site. The analysis is based on the proposition that if a particular explanatory variable governs a response variable across a large spatial extent, then one would expect that the degree of synchrony in the response variable between two sites (measured by temporal cross-correlation) should be largely determined by the synchrony in the explanatory variable between the same sites. When both variables are measured at several sites, one would expect concordance (a positive correlation) between synchrony of a causal factor and synchrony of the response. Therefore,

we constructed a separate matrix of pairwise synchrony estimates (cross-correlations between sites, transformed by Hotelling's z^* transformation; Sokal and Rohlf 1981) for each variable (detrended LDI, PHDI, MST, MWT, and TSP). We then measured spatiotemporal concordance by the Pearson correlation between the LDI matrix and each weather matrix (Schauber et al. 2002). The significance of each concordance measure was evaluated by Mantel randomization tests with 10 000 replications (Manly 1997). Statistically significant spatiotemporal concordance could result from spatial autocorrelation in both variables, so this procedure represents a liberal test (Type I error rate higher than nominal) of hypothesized predictors.

RESULTS

LDI increased substantially between 1990 and 2002 in Dutchess County and in each state (Fig. 2). Based on AIC_c values, the most parsimonious trend model was quadratic for Dutchess County, Massachusetts, Rhode Island, and Pennsylvania (Fig. 2A, C, E, F); linear trend models were more parsimonious for Connecticut, Maryland, New Jersey, and New York (Fig. 2B, D, G, H).

Analysis of each predictor individually

For the period 1994-2002, when data were available for all explanatory variables, detrended LDI was most strongly correlated with either acorn production or mouse abundance in Dutchess County, Rhode Island, Connecticut, and Pennsylvania (Table 1). Correlations of LDI with acorns or mice were statistically significant for Dutchess County, New York State, and Connecticut. PHDI2 was the strongest predictor for Massachusetts, and had moderate to high correlations (r > 0.3) with detrended LDI for all sites except New Jersey and Maryland. After adjusting for autocorrelation, however, correlations with PHDI2 were statistically significant only for Massachusetts and Connecticut. Observed correlations between TSP1 and LDI were negative at all sites, significantly so for New York and Connecticut, and TSP1 was the strongest predictor for New York. MST1 showed positive correlations with LDI at all sites, with moderate (0.3 < r < 0.6) correlations for Dutchess County (P < 0.05), New York State, New Jersey, and Maryland. Correlations between MWT1 and LDI were highly variable ($-0.33 \le r \le$

Although mouse abundance tended to be a good predictor of LDI over the years 1994–2002, expanding the analysis to include 1992 and 1993 led to substantially lower correlations. Mouse abundance was moderate in 1991 and very high in 1992, whereas detrended LDI in the following years showed the opposite pattern at several sites. For 1992–2002, PHDI2 was the strongest predictor for all sites except New York State and New Jersey, and had moderate to high correlations with LD incidence for all sites except New Jersey (Table 1).

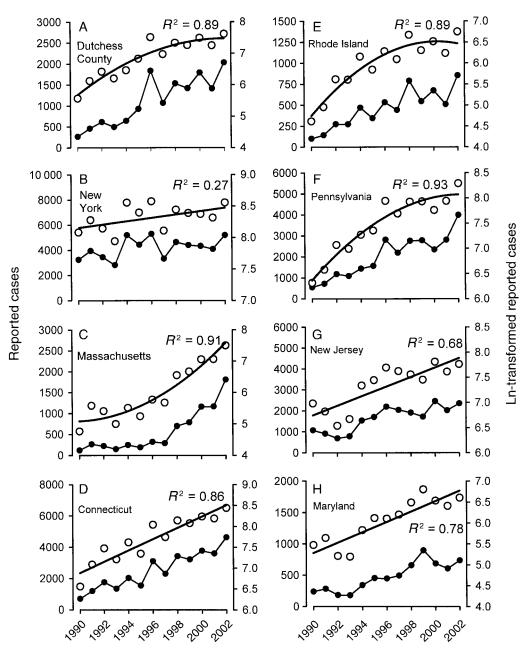


Fig. 2. Long-term trends in reported Lyme disease incidence before (solid circles) and after In-transformation (open circles). Results are shown for (A) Dutchess County, New York; (B) New York State, (C) Massachusetts, (D) Connecticut, (E) Rhode Island, (F) Pennsylvania, (G) New Jersey, and (H) Maryland. Trend lines are linear or quadratic regressions against year minus 1989.

TSP1 again showed moderate negative correlations with LDI for all sites except Maryland, but the correlation was statistically significant only for Connecticut. Correlations between MST1 and LDI were positive at all sites, and MST1 was a significant predictor for New York State and Dutchess County. As was the case for the 1994–2002 period, correlations with MWT1 were inconsistent, ranging from -0.43 to 0.76. However, MWT1 was a significant positive predictor for Massachusetts and Connecticut.

Multiple-predictor model selection

The most parsimonious model for nearly every site included ≤ 1 explanatory variable (Table 2). The few exceptions involved models including two weather variables. All models with three or more predictor variables performed poorly ($R^2 < 0.1$ or δ -AIC $_c > 3.0$). The null model was a strong candidate for the best approximating model (δ -AIC $_c < 1.5$) for Rhode Island, Pennsylvania, New Jersey, and Maryland, indicating little support for the hypothesis that LDI fluctuations

Table 1. Cross-correlations between detrended Lyme disease incidence and hypothesized predictors for Dutchess County, New York (DC), and seven northeastern U.S. states.

		State							
Predictor	Years (t)	DC	NY	MA	CT	RI	PA	NJ	MD
Acorns $t-2$ Mice $t-1$ PHDI $t-2$ Winter temp. $t-1$ Summer temp. $t-1$ Summer precip. $t-1$	1994–2002 1994–2002 1994–2002 1994–2002 1994–2002 1994–2002	0.87 0.78 0.63 0.25 0.57 -0.55	0.31 0.62 0.52 0.00 0.53 -0.68	0.29 0.46 0.85 0.82 0.12 -0.43	0.56 0.81 0.58 0.50 0.17 - 0.79	0.08 0.45 0.39 0.11 0.25 -0.31	0.36 0.56 0.49 -0.03 0.03 -0.26	0.51 0.43 0.04 -0.33 0.53 -0.29	0.21 0.09 0.28 0.33 0.38 -0.17
Mice $t-1$ PHDI $t-2$ Winter temp. $t-1$ Summer temp. $t-1$ Summer precip. $t-1$	1992–2002 1992–2002 1992–2002 1992–2002 1992–2002	0.46 0.71 0.27 0.68 -0.57	0.25 0.55 -0.08 0.59 -0.60	0.19 0.88 0.76 0.39 -0.52	0.52 0.65 0.52 0.34 -0.64	0.35 0.44 0.19 0.21 -0.28	0.36 0.51 0.02 0.23 -0.35	0.06 -0.05 -0.43 0.43 -0.33	-0.18 0.35 -0.01 0.29 0.02

Note: Values in boldface were significant (one-sided P < 0.05) after adjusting for serial autocorrelation.

in these states are predictable on the basis of any of the variables that we examined. Massachusetts LDI data exhibited no evidence of predictability on the basis of acorns or mice measured at IES, but showed high coefficients of determination ($R^2 \geq 0.72$) based on PHDI2, MWT1, or a combination of both. For the years 1994–2002, mice and TSP1 each performed well in predicting LDI in Connecticut and New York State, and acorn production was the sole supported variable for Dutchess County. However, for 1992–2002, mouse abundance was a poor predictor for all states, whereas PHDI2 was in the best models for Dutchess County, Massachusetts, Connecticut, Rhode Island, and Pennsylvania.

Spatiotemporal concordance

From 1990 to 2002, interstate synchrony was high for all weather variables (Fig. 1B–E and Fig. 3). Fluctuations in detrended LDI showed moderate synchrony (r > 0.4) among Connecticut, Massachusetts, and New York, but weak synchrony (r < 0.25) between Pennsylvania and Massachusetts and between New Jersey and Maryland and all other states except New York (Fig. 3). Synchrony in LDI showed essentially zero concordance with synchrony of PHDI and MWT, and concordance with MST and TSP was nonsignificant.

DISCUSSION

Understanding and predicting temporal fluctuations in Lyme disease risk would be beneficial in allowing public health officials to focus intervention efforts during years of greatest risk. Several biotic and abiotic variables have been proposed as potential predictors, but they have not been rigorously confronted with a common data set. Our goal was to objectively assess the ability of weather variables, acorn production, and mouse abundance to predict state or local (county) reports of Lyme disease incidence (LDI). We also sought to provide more rigorous statistical analyses than have been done previously, by explicitly accounting for the temporal autocorrelation among successive annual

measures of LDI, and by using information theoretic approaches to model selection (Burnham and Anderson 1998).

Our results did not indicate that any one variable or category of variables consistently outperformed others, and LDI in four out of the seven states (Rhode Island, New Jersey, Pennsylvania, and Maryland) showed no evidence of predictability on the basis of the variables that we examined. Acorn production and mouse abundance showed greatest predictive power for Dutchess County, the only sites where these variables were measured, but also tended to be positively correlated with Lyme disease incidence for the other states that we examined, except Maryland. After accounting for serial autocorrelation, PHDI two years previously was a significant predictor for two states and was essentially uncorrelated with Lyme incidence in New Jersey. For the sites that we examined, there was a consistent negative relationship between total precipitation in the previous summer and Lyme disease incidence. Prior-year summer temperatures showed consistent positive correlation with Lyme disease incidence, but the correlation was only strong and statistically signficant for New York. Correlations between prior-year winter temperature and Lyme disease incidence were inconsistent. Information theoretic model selection suggested that models with more than two variables were not parsimonious, so combining predictors did not substantially improve predictive performance.

Our results highlight the difficulties in extracting rigorous causal inference from short-term (13 years) observational data subject to spatial and temporal autocorrelation. For example, detrended Lyme disease incidence and weather variables exhibited synchrony across the seven states included in this study. Therefore, finding that a weather variable had predictive power for several states does not actually provide much additional independent information. If one of the weather variables that we examined truly governs fluctuations in Lyme disease incidence across our study region, then states with similar fluctuations in weather

TABLE 2. Results of best subset model selection for predicting detrended Lyme disease incidence in Dutchess County, New York, and seven northeastern U.S. states.

Time period and site	Variables included	δ-AIC _c	R^2
1994-2002			
Dutchess County	Acorns $t-2$	0	0.75
New York State	TSP1 Mice $t - 1$ MST1 Null	0 1.1 2.5 2.7	0.46 0.39 0.28
Massachusetts	PHDI2 PHDI2, MWT1 MWT1	0 0.4 1.5	0.72 0.87 0.67
Connecticut	Mice $t - 1$ TSP1 MST1, TSP1	0 0.6 2.4	0.65 0.63 0.78
Rhode Island	Null	0	
Pennsylvania	Mice $t - 1$ Null	0 0.6	0.32
New Jersey	MST1 Null	$0 \\ 0.2$	0.28
Maryland	Null	0	
1992-2002			
Dutchess County	PHDI2 MST1 PHDI2, MST1	0 0.9 2.2	0.50 0.46 0.62
New York State	TSP1 MST1 PHDI2 MWT1, TSP1 PHDI2, MWT1 Null	0 0.06 0.9 1.6 2.5 2.9	0.35 0.35 0.30 0.54 0.50
Massachusetts	PHDI2, MWT1 PHDI2	$0 \\ 0.8$	$0.87 \\ 0.77$
Connecticut	PHDI2 TSP1 MWT1, TSP1 Mice t – 1 MWT1 PHDI2, TSP1	0 0.2 2.4 2.5 2.5 2.7	0.42 0.41 0.55 0.27 0.27 0.54
Rhode Island	PHDI2 Null	0 0.4	0.19
Pennsylvania	PHDI2 Null	0 1.4	0.26
New Jersey	MWT1, TSP1 MWT1 MST1 Null	0 0.4 0.5 0.7	0.51 0.19 0.18
Maryland	Null	0	

Note: For each time period and site, we list variables included in the best model (δ -AIC_c = 0) and models meeting all the following criteria: δ -AIC_c < 3.0, δ -AIC_c < δ -AIC_c[Null], and $R^2 > 0.1$.

should exhibit similar fluctuations in Lyme incidence. However, we were unable to demonstrate spatiotemporal concordance between Lyme incidence and any weather variable. This suggests that either the variables that we examined have little true effect on LDI, or that different factors predominate in different areas. The latter possibility casts doubt on the hope that a simple,

general explanation for LDI fluctuations will be uncovered. Future observations will be needed to confirm or refute the hypothesis that the weather variables we examined are causally related to subsequent fluctuations in LDI.

Weather-related variables are logical candidates to explain and predict future fluctuations in LDI, but many links in hypothesized causal chains have not been validated empirically. For example, Subak (2003) speculated that the observed relationship between PHDI and Lyme risk two years later might stem from drought reducing the survival of nymphal ticks, thereby reducing the abundance of adult ticks in that year. Reduced adult abundance would then result in reduced abundance of infected nymphs two years later. Of these steps, only the effect of moisture on nymphal survival is supported by empirical field studies (Lindsay et al. 1998), and it is brought into doubt by the inconsistent effect of PHDI on concurrent Lyme incidence (Subak 2003). Although high nymphal mortality in one generation logically might be expected to reduce nymphal abundance in the next generation, this has not been demonstrated, and nymphal mortality could be compensated for by increases in survival or reproductive success of intervening life stages.

Unlike weather-related hypotheses, the hypothesis that acorn production indirectly affects Lyme disease risk is buttressed by a logical chain of causation with empirical support for each link: from acorns to smallmammal abundance (Elkinton et al. 1996, Wolff 1996, Jones et al. 1998), from acorns to the abundance and distribution of larval ticks (Jones et al. 1998), and ultimately to the abundance of infected nymphal ticks (Ostfeld et al. 2001). Here, we complete the chain by empirically linking past acorn production with observed fluctuations in LDI in humans, locally and perhaps regionally. Clearly, Lyme disease occurs in landscapes that are not dominated by oaks (Ginsberg et al. 1998, Ostfeld et al. 1998), and in these areas we would expect mouse populations and LDI to fluctuate independently of acorn crops. The observation that acorns and mice measured in Dutchess County can predict LDI fluctuations in other states as well as or better than weather variables measured in those states suggests that local impacts of acorns and their mammalian consumers are potent, but that broad-scale impacts are diluted by other factors that predominate in non-oak-dominated landscapes.

Spatial synchrony in acorn production dictates the spatial extent over which acorn data from one site could potentially predict LDI. Large-scale synchrony in acorn production has been documented (Downs and McQuilkin 1944, Koenig and Knops 2000). Unfortunately, long time series of acorn production and mouse abundance are much more difficult to obtain than similar weather data, and we do not know to what degree acorn production is synchronized among northeastern states. However, some short-term comparisons of acorn pro-

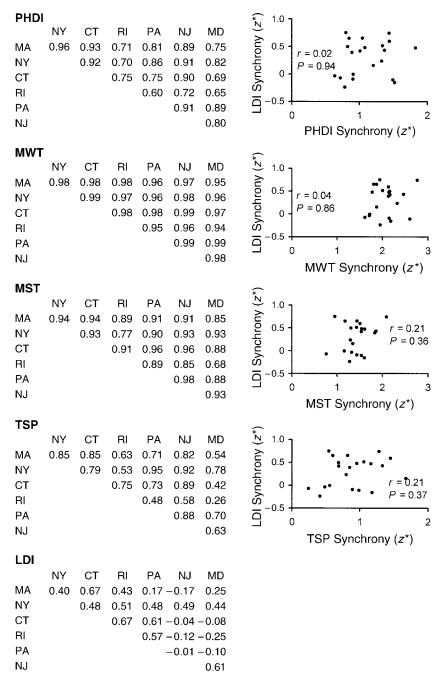


Fig. 3. Matrices of synchrony between states for each weather variable and Lyme disease incidence (LDI), and plots of LDI synchrony (after Hotelling's z^* transformation) against weather variable synchrony. Correlations reported on scatterplots are measures of spatiotemporal concordance, and P values result from Mantel randomization tests. Abbreviations: PDHI, Palmer Hydrologic Drought Index; MWT, mean winter temperature; MST, mean summer temperature; TSP, total summer precipitation.

duction can be made. Healy et al. (1999) published a time series from 1986 to 1996 of acorn production by red oaks (*Quercus rubra*) in western Massachusetts. Their measurements of acorn production were broadly consistent with our measurements at IES, in that acorn production was near zero in 1992, moderate or high in 1993–1994, low in 1995, and moderate in 1996. Visual

inspection of their data suggests a negative relationship with detrended Lyme disease incidence in Massachusetts two years later. However, most cases of Lyme disease in Massachusetts come from the eastern portion of the state. Data published by Elkinton et al. (1989) indicate that fluctuations in acorns and mice in eastern and western Massachusetts may be asynchronous.

Therefore, predicting Lyme disease risk in eastern Massachusetts might require local acorn measurements. Liebhold et al. (2000) published measurements of acorn production in central Pennsylvania from 1968 to 1994. These data were broadly concordant with acorn production at IES, with acorn failure in 1992 and increasing production in 1993–1994. Furthermore, large acorn crops reported for 1990 and 1994 were followed by high values of detrended Lyme disease incidence in 1992 and 1996 in Pennsylvania.

Subak (2003) implied that fluctuations in winter temperatures may be a more important factor in mouse population dynamics than acorn production. However, this inference was based on analysis of a six-year time series of mouse abundance at IES. Our full data set indicates that mouse abundance from 1993 to 2002 was much more strongly correlated with prior acorn production (r = 0.81) than with prior winter temperatures (r = 0.49). Subak (2003) also suggested that even a temporary effect of winter temperatures on mouse abundance may be important for the success of larval ticks hatching in spring. However, most larval blacklegged ticks in the northeastern United States hatch in late summer (Fish 1993, Ostfeld et al. 1996a, b), so this mechanism seems dubious. In perhaps the most comprehensive study of P. leucopus population dynamics, Lewellen and Vessey (1998) found that the response of mouse numbers to weather lasted only 1-2 months, and that summer peak densities (most relevant to Lyme disease epidemiology) were unaffected by weather, with the exception of two droughts over the 23-year study. The weak effect of winter temperatures on mouse abundance may explain the inconsistent correlations between winter temperatures and Lyme disease incidence.

We were somewhat surprised by the consistent negative relationship between summer precipitation and Lyme disease incidence in the next year. This result implies that wet summers might actually prove detrimental to immature blacklegged ticks, perhaps by enhancing the abundance or efficacy of natural enemies. Entomopathogenic soil fungi (Metarhizium and Beauveria), which seem capable of controlling blacklegged tick populations under some conditions (Samish and Rehacek 1999, Benjamin et al. 2002), might benefit from high summer precipitation and subsequently might reduce tick abundance. Recently, Chase and Knight (2003) found that wet conditions supported predators and competitors capable of reducing wetland mosquito populations, with the surprising result that mosquito populations were more abundant following dry than wet years. Whether the same might be true for ticks is not known.

Our results do not present a clear answer to the question of whether acorns, mice, or any of the weather variables that we examined is a reliable predictor of Lyme disease incidence over large areas of the northeastern United States. However, confronting a priori

predictions against future Lyme incidence levels may provide critical tests of competing hypotheses. Acorn production was relatively high at IES in 2001 and 2002 (Fig. 1A), implying that Lyme disease incidence will be greater than the trend would suggest for 2003 and 2004. PHDI was strongly negative in most of the states in our study in 2002 (Fig. 1B), and total summer precipitation was high in 2003 (Fig. 1C). These two weather factors both point to lower than expected Lyme disease incidence in 2004. Finally, our analyses did not account for potential biennial dynamics in tick abundance. Because *I. scapularis* has a two-year life cycle, biennial dynamics (Rost et al. 2001) also could explain fluctuations in Lyme disease incidence. Detrended LDI appears to alternate between positive and negative values in several states (Fig. 2), which may reflect biennial dynamics of tick abundance; however, summer precipitation exhibits a similar pattern (Fig. 1C). The twovear pattern in summer precipitation was broken in 2002, however. If biennial dynamics predominate, we would expect lower incidence in 2003 and higher in 2004 for most of the states in our study. Thus, the acorn cascade, weather, and biennial dynamics hypotheses make conflicting forecasts regarding LDI in the next two years, presenting an opportunity for a critical test of these competing hypotheses.

Analyses of trends in numbers of Lyme disease cases reported by state health departments to the CDC are subject to temporal and spatial variability in criteria for reporting and, consequently, in the likelihood that the data represent true numbers of Lyme disease cases. Under-reporting can occur when awareness of Lyme disease by at-risk citizens and their healthcare providers is low (such as when it initially enters a geographic area), and when local health departments lack funds to identify and record cases. If awareness and funding stabilize at sufficiently high levels, these sources of error are likely to be reduced in the future.

Our objective was to compare the performance of previously identified variables in predicting Lyme disease incidence, and not to search for new predictors. It is possible that a different weather variable may prove to be a reliable predictor, but demonstrating so will require more exploratory analyses confirmed by rigorous tests against independent data sets. At this point, it remains premature to claim that any biotic or abiotic variable provides superior ability to predict Lyme disease incidence. Longer time series of response and predictor variables and standardized case reporting will aid in discriminating among purported predictors, and future exploratory analyses may identify more reliable predictors. Realistically, it seems likely that both biotic and abiotic factors contribute to observed fluctuations in Lyme disease incidence, so incorporating multiple variables may ultimately be necessary to achieve reliable prediction.

ACKNOWLEDGMENTS

This research was supported by grants from the Andrew W. Mellon Foundation, the National Science Foundation

(DEB-9807115, DEB-0075277), and the National Institutes of Health (R01AI40076). The authors would like to thank F. Keesing, C. Jones, K. LoGiudice, K. Schmidt, B. Goodwin, and M. Eichholz for stimulating discussions and review of previous drafts. Constructive criticism from two anonymous reviewers substantially improved the manuscript. This paper is a contribution to the program of the Institute of Ecosystem Studies.

LITERATURE CITED

- Abraham, B., and J. Ledolter. 1983. Statistical methods for forecasting. John Wiley, New York, New York, USA.
- Anderson, D. R., K. P. Burnham, and W. L. Thompson. 2000. Null hypothesis testing: problems, prevalence, and an alternative. Journal of Wildlife Management 64:912–923.
- Arnason, A. N., and C. J. Schwartz. 1999. Using POPAN-5 to analyze banding data. Bird Study 46(Supplement):S127–S168.
- Benjamin, M. A., E. Zhioua, and R. S. Ostfeld. 2002. Laboratory and field evaluation of the entomopathogenic fungus *Metarhizium anisopliae* (Deuteromycetes) for controlling questing adult *Ixodes scapularis* (Acari: Ixodidae). Journal of Medical Entomology **39**:723–728.
- Bjørnstad, O. N., R. A. Ims, and X. Lambin. 1999. Spatial population dynamics: analyzing patterns and processes of population synchrony. Trends in Ecology and Evolution 14: 427–432.
- Burnham, K. P., and D. R. Anderson. 1998. Model selection and inference: a practical information-theoretic approach. Springer-Verlag, New York, New York, USA.
- Centers for Disease Control and Prevention. 1997. Case definitions for infectious conditions under public health surveillance. Morbidity and Mortality Weekly Report **46**:1–55.
- Centers for Disease Control and Prevention. 2002. Lyme Disease—United States, 2000. Morbidity and Mortality Weekly Report **51**:29–31.
- Centers for Disease Control and Prevention. 2003. Morbidity and Mortality Weekly Report 51:1169.
- Chase, J. M., and T. M. Knight. 2003. Drought-induced mosquito outbreaks in wetlands. Ecology Letters 6:1017–1024.
- Chow, C., E. B. Hayes, A. S. Evans, G. Johnson, S. Marks, M. Caldwell, D. White, and C. Noonan-Toly. 2002. Surveillance for Lyme disease—Dutchess County, New York, 1992–2000. National Center for Infectious Diseases, Atlanta, Georgia, USA.
- Donahue, J. G., J. Piesman, and A. Spielman. 1987. Reservoir competence of white-footed mice for Lyme disease spirochetes. American Journal of Tropical Medicine and Hygiene **36**:92–96.
- Downs, A. A., and W. E. McQuilkin. 1944. Seed production of southern Appalachian oaks. Journal of Forestry 42:913– 920.
- Elkinton, J. S., J. R. Gould, A. M. Liebhold, H. R. Smith, and W. E. Wallner. 1989. Are gypsy moth populations in North America regulated at low density? Pages 233–249 in W. E. Wallner and K. A. McManus, editors. Lymantridae: a comparison of features of New and Old World tussock moths. USDA Forest Service, Northeastern Forest Experiment Station, New Haven, Connecticut, USA.
- Elkinton, J. S., W. M. Healy, J. P. Buonaccorsi, G. H. Boettner, A. M. Hazzard, H. R. Smith, and A. M. Liebhold. 1996. Interactions among gypsy moths, white-footed mice, and acorns. Ecology 77:2332–2342.
- Falco, R. C., and D. Fish. 1989. Potential for tick exposures in recreational parks in a Lyme disease endemic area. American Journal of Public Health 79:12–15.
- Falco, R. C., D. F. McKenna, T. J. Daniels, R. B. Nadelman, J. Nowakowski, D. Fish, and G. P. Wormser. 1999. Temporal relation between *Ixodes scapularis* abundance and

- risk for Lyme disease associated with erythema migrans. American Journal of Epidemiology **149**:771–776.
- Fish, D. 1993. Population ecology of *Ixodes dammini*. Pages 25–42 *in* H. S. Ginsberg, editor. Ecology and management of Lyme disease. Rutgers University Press, Piscataway, New Jersey, USA.
- Ginsberg, H. S., K. E. Hyland, R. Hu, T. J. Daniels, and R. C. Falco. 1998. Tick population trends and forest type. Science 281:350–351.
- Healy, W. M., A. M. Lewis, and E. F. Boose. 1999. Variation of red oak acorn production. Forest Ecology and Management 116:1–11.
- Jones, C. G., R. S. Ostfeld, M. P. Richard, E. M. Schauber, and J. O. Wolff. 1998. Chain reactions linking acorns to gypsy moth outbreaks and Lyme disease risk. Science 279: 1023–1026.
- Jones, C. J., and U. D. Kitron. 2000. Populations of *Ixodes scapularis* (Acari: Ixodidae) are modulated by drought at a Lyme disease focus in Illinois. Journal of Medical Entomology 37:408–415.
- Koenig, W. D. 1999. Spatial autocorrelation in ecological studies. Trends in Ecology and Evolution 14:22–26.
- Koenig, W. D., and J. M. H. Knops. 2000. Patterns of annual seed production by Northern Hemisphere trees: a global perspective. American Naturalist **155**:59–69.
- Lewellen, R. H., and S. H. Vessey. 1998. The effect of density dependence and weather on population size of a polyvoltine species. Ecological Monographs 68:571–594.
- Liebhold, A., J. Elkinton, D. Williams, and R. Muzika. 2000. What causes outbreaks of the gypsy moth in North America? Population Ecology 42:257–266.
- Liebhold, A., V. Sork, M. Peltonen, W. Koenig, O. N. Bjørnstad, R. Westfall, J. Elkinton, and J. M. H. Knops. 2004. Within-population spatial synchrony in mast seeding of North American oaks. Oikos 104:156–164.
- Lindsay, L. R., I. K. Barker, G. A. Surgeoner, S. A. McEwen, T. J. Gillespie, and E. M. Addison. 1998. Survival and development of the different life stages of *Ixodes scapularis* (Acari: Ixodidae) held within four habitats on Long Point, Ontario, Canada. Journal of Medical Entomology 35:189– 199.
- Lindsay, L. R., S. W. Mathison, I. K. Barker, S. A. McEwen, T. J. Gillespie, and G. A. Surgeoner. 1999. Microclimate and habitat in relation to *Ixodes scapularis* (Acari: Ixodidae) populations on Long Point, Ontario, Canada. Journal of Medical Entomology 36:255–262.
- LoGiudice, K., R. S. Ostfeld, K. A. Schmidt, and F. Keesing. 2003. The ecology of infectious disease: effects of host diversity and community composition on Lyme disease risk. Proceedings of the National Academy of Sciences (USA) 100:567-571.
- Manly, B. F. J. 1997. Randomization, bootstrap and Monte Carlo methods in biology. Second edition. Chapman and Hall, London, UK.
- Mather, T. N., M. C. Nicholson, E. F. Donnelly, and B. T. Matyas. 1996. Entomologic index for human risk of Lyme disease. American Journal of Epidemiology 144:1066–1069.
- Mather, T. N., M. L. Wilson, S. I. Moore, J. M. C. Ribeiro, and A. Spielman. 1989. Comparing the relative potential of rodents as reservoirs of the Lyme disease spirochete (*Borrelia burgdorferi*). American Journal of Epidemiology 130:143–150.
- McShea, W. J., and G. Schwede. 1993. Variable acorn crops: responses of white-tailed deer and other mast consumers. Journal of Mammalogy **74**:999–1006.
- Meek, J. I., C. L. Roberts, E. V. Smith, Jr., and M. L. Cartter. 1996. Underreporting of Lyme disease by Connecticut physicians. Journal of Public Health Management Practice 2: 61–65.

- New York State Department of Health. 2003. Communicable diseases in New York State: annual reports. http://www.health.state.ny.us/nysdoh/epi/mainrpt.htm
- Orloski, K. A., E. B. Hayes, G. L. Campbell, and D. T. Dennis. 2000. Surveillance for Lyme disease—United States, 1992–1998. CDC Surveillance Summaries. Monthly Morbidity and Mortality Weekly Report 49:1–11.
- Ostfeld, R. S. 1997. The ecology of Lyme disease risk. American Scientist **85**:338–346.
- Ostfeld, R. S., O. M. Cepeda, K. R. Hazler, and M. C. Miller. 1995. Ecology of Lyme disease: habitat associations of ticks (*Ixodes scapularis*) in a rural landscape. Ecological Applications 5:353–361.
- Ostfeld, R. S., K. R. Hazler, and O. M. Cepeda. 1996a. Temporal and spatial dynamics of *Ixodes scapularis* (Acari: Ixodidae) in a rural landscape. Journal of Medical Entomology **33**:90–95.
- Ostfeld, R. S., C. G. Jones, M. P. Richard, E. M. Schauber, and J. O. Wolff. 1998. Tick population trends and forest type. Science **281**:350–351.
- Ostfeld, R. S., C. G. Jones, and J. O. Wolff. 1996. Of mice and mast. Bioscience 46:323-330.
- Ostfeld, R. S., M. C. Miller, and K. R. Hazler. 1996b. Causes and consequences of tick (*Ixodes scapularis*) burdens on white-footed mice (*Peromyscus leucopus*). Journal of Mammalogy **77**:266–273.
- Ostfeld, R. S., E. M. Schauber, C. D. Canham, F. Keesing, C. G. Jones, and J. O. Wolff. 2001. Effects of acorn production and mouse abundance on abundance and *Borrelia burgdorferi* infection prevalence of nymphal *Ixodes scapularis* ticks. Vector Borne and Zoonotic Diseases 1:55–63.
- Rost, M., G. Várkonyi, and I. Hanski. 2001. Patterns of 2-year population cycles in spatially extended host-parasitoid systems. Theoretical Population Biology **59**:223–233.
- Samish, M., and J. Rehacek. 1999. Pathogens and predators of ticks and their potential in biological control. Annual Review of Entomology 44:159–182.
- Schauber, E. M., D. Kelly, P. Turchin, C. Simon, W. G. Lee, R. B. Allen, I. J. Payton, P. R. Wilson, P. E. Cowan, and

- R. E. Brockie. 2002. Masting by 18 New Zealand plant species: the role of temperature as a synchronizing cue. Ecology **83**:1214–1225.
- Schmidt, K. A., R. S. Ostfeld, and E. M. Schauber. 1999. Infestation of *Peromyscus leucopus* and *Tamias striatus* by *Ixodes scapularis* (Acari: Ixodidae) in relation to the abundance of hosts and parasites. Journal of Medical Entomology **36**:749–757.
- Sciremammano, F., Jr. 1979. A suggestion for the presentation of correlations and their significance levels. Journal of Physical Oceanography 9:1273–1276.
- Slajchert, T., U. D. Kitron, C. J. Jones, and A. Mannelli. 1997. Role of the eastern chipmunk (*Tamias striatus*) in the epizootiology of Lyme borreliosis in northwestern Illinois, USA. Journal of Wildlife Diseases 33:40–46.
- Sokal, R. R., and F. J. Rohlf. 1981. Biometry: the principles and practice of statistics in biological research. Second edition. W. H. Freeman, New York, New York, USA.
- Stafford, K. C., III. 1994. Survival of immature *Ixodes sca-pularis* (Acari: Ixodidae) at different relative humidities. Journal of Medical Entomology **31**:310–314.
- Stafford, K. C., III, M. L. Cartter, L. A. Magnarelli, S. H. Ertel, and P. A. Mshar. 1998. Temporal correlations between tick abundance and prevalence of ticks infected with *Borrelia burgdorferi* and increasing incidence of Lyme disease. Journal of Clinical Microbiology 36:1240–1244.
- Subak, S. 2003. Effects of climate on variability in Lyme disease incidence in the Northeastern United States. American Journal of Epidemiology 157:531–538.
- Vail, S. G., and G. Smith. 2002. Vertical movement and posture of blacklegged tick (Acari: Ixodidae) nymphs as a function of temperature and relative humidity in laboratory experiments. Journal of Medical Entomology 39:842–846.
- Wolff, J. O. 1996. Population fluctuations of mast-eating rodents are correlated with production of acorns. Journal of Mammalogy 77:850–856.
- Yoder, J. A., and A. Spielman. 1992. Differential capacity of larval deer ticks (*Ixodes dammini*) to imbibe water from subsaturated air. Journal of Insect Physiology 38:863–869.