



Practice of Epidemiology

Predicting Human West Nile Virus Infections With Mosquito Surveillance Data

A. Marm Kilpatrick* and W. John Pape

* Correspondence to Dr. A. Marm Kilpatrick, Department of Ecology and Evolutionary Biology, 1156 High Street, University of California, Santa Cruz, Santa Cruz, CA 95064 (e-mail: akilpatr@ucsc.edu).

Initially submitted September 18, 2012; accepted for publication February 21, 2013.

West Nile virus (WNV) has become established across the Americas with recent heightened activity causing significant human illness. Surveillance methods to predict the risk of human infection are urgently needed to initiate timely preventative measures and justify the expense of implementing costly or unpopular control measures, such as aerial spraying or curfews. We quantified the links between mosquito surveillance data and the spatiotemporal patterns of 3,827 human WNV cases reported over 5 years in Colorado from 2003 to 2007. Mosquito data were strongly predictive of variation in the number of human WNV infections several weeks in advance in both a spatiotemporal statewide analysis and temporal variation within counties with substantial numbers of human cases. We outline several ways to further improve the predictive power of these data and we quantify the loss of information if no funds are available for testing mosquitoes for WNV. These results demonstrate that mosquito surveillance provides a valuable public health tool for assessing the risk of human arboviral infections, allocating limited public health resources, and justifying emergency control actions.

arbovirus; disease control; eastern equine encephalitis virus; Lyme disease; predictive model; public health; vector index

Abbreviations: CDC, Centers for Disease Control and Prevention; WNV, West Nile virus.

The surveillance of enzootic transmission of zoonotic pathogens is critical to reducing human disease and is necessary to distinguish between annual enzootic circulation and periodic epizootic transmission (1, 2). Surveillance can be used to determine whether a pathogen is present and to assess the intensity of transmission and the risk for human infection (3). If accurate estimates of an increasing risk of human transmission can be forecast in advance, public health and control resources can be activated and directed to the areas of highest risk. Conversely, in years with low risk of epidemic transmission, preventative measures can be scaled back appropriately to conserve limited funds.

Surveillance of vector-borne zoonotic pathogens, which are transmitted between vectors and human and nonhuman vertebrates, can focus on the intensity of transmission in wild-life hosts, vectors, or humans (4–6). Key challenges are identifying which methods to use for surveillance, how to accurately interpret surveillance findings with respect to human risk, and how to present these findings to elected officials and the public

in a meaningful context (3). West Nile virus (WNV), which is transmitted primarily between mosquitoes and birds but also infects the human population, is emblematic of these surveillance challenges.

The emergence of WNV in New York State in 1999 and the significant morbidity and mortality it caused as it spread rapidly throughout the Americas have been well documented (7–9). Reported WNV cases in North America through December 11, 2012, have included 16,122 cases of encephalitis (41,493 total cases) and 1,594 deaths that resulted from an estimated 2.3 million infections (7, 10, 11). In a substantial fraction of these cases, there are long-term sequelae or lasting illness, including partial paralysis, chronic fatigue, and other rarer symptoms (12).

The following 4 methods of WNV animal surveillance are described in the Centers for Disease Control and Prevention (CDC) publication, *Guidelines for Surveillance, Prevention, and Control of West Nile Virus Infection, 2000*, and subsequent updates (3, 8, 13): dead bird reporting and testing,

mosquito trapping and testing, monitoring the serostatus of sentinel chicken flocks, and monitoring of equine infections as sentinels of mammalian WNV transmission. A key challenge is to recommend “trigger points” based on animal and human case surveillance for initiating public health and mosquito control actions.

Several challenges and limitations in data interpretation have become apparent with 3 of the 4 surveillance methods. First, although dead bird testing and reporting were effectively used to monitor the spread of WNV across the United States (10, 14), converting data from dead bird reports into a risk index requires large numbers of dead bird reports, and the mathematical tools required to analyze the data are challenging for local public health agencies to use (15). In addition, declining public interest in WNV, decreased corvid densities in many areas due to direct WNV mortality (16, 17), and the labor-intensive work required by local agencies to collect and submit dead birds have decreased the utility of this surveillance technique (18). In Colorado, for example, the number of bird carcasses tested dropped from 889 and 1,575 in 2002 and 2003, respectively, to 42 in 2010 (Colorado Department Public Health and Environment, unpublished data). Second, although sentinel chicken flocks have been shown to be effective in some areas, they were ineffective for WNV detection in many others, including Colorado, and were eliminated from the Colorado state surveillance program in 2004 (19). Finally, although many health departments passively collect results of equine antibody testing from veterinary diagnostic laboratories, equine cases decreased greatly because of immunity from natural infections and equine WNV vaccines and an increasing number of veterinarians making clinical diagnoses without submitting serum samples. Thus, these 3 surveillance methods have severe limitations in their ability to quantify the risk of WNV transmission to humans in a timely and meaningful manner.

The remaining method, mosquito surveillance, offers several advantages, allowing rapid measurements of vector mosquito population densities and infection rates and determinations of whether these are stable, increasing, or decreasing. Additionally, integration of mosquito abundance and infection data into what has been called a “vector index” that measures the abundance of WNV-infected mosquitoes (the entomological risk) should theoretically provide a quantitative measure of the risk of human WNV infection (20). If a predictive relationship can be established and a threshold value can be determined that signals impending human risk, this could be readily conveyed and understood by the public and elected officials when deciding whether to undertake WNV prevention efforts. Elected officials frequently request a trigger point, or threshold, above which they can report to their constituents that a specified level of human risk has been reached, and the actions being taken and expenditures being made are justified. Thus, a transparent and simple risk index is often useful to gain support for public health recommendations.

Although many health agencies collect mosquito surveillance data, including *Culex* population abundance and minimum infection rates, and some have integrated these into a “vector index” that is used as a trigger point for action, there has been little validation of the correlation between mosquito

surveillance data and the spatiotemporal risk of human WNV infection. The key challenge is to determine the power of mosquito surveillance data to predict the number of human cases far enough in advance to allow the initiation of timely, preventative measures and to develop thresholds for action. Four studies have attempted to assess the predictive power of mosquito surveillance data for the risk of human WNV infection. Two studies examined temporal variation in a vector index (the product of the abundance and WNV prevalence of a single mosquito species) and human WNV cases in a single region and found they were highly correlated (18, 21). A third study used an environmental model to predict spatial variation in mosquito abundance and then used this to generate a categorical index of mosquito abundance to examine spatial variation in human WNV infection within and between counties (22). Positive associations existed between categorical estimates of mosquito abundance and the risk of human infection in some regions, but unexpected negative relationships occurred in others. The fourth study, in the mid-Atlantic region of the United States (23), integrated the following 4 aspects of mosquito ecology and epidemiology into a risk index: mosquito abundance, WNV infection prevalence, vector competence, and mosquito feeding patterns (20). This index estimates the abundance of WNV-transmitting mosquitoes that will feed on humans and was highly correlated with temporal variation in human WNV cases in the mid-Atlantic region, with 60% of the variance explained (23). One advantage of this risk index is that it translates mosquito surveillance data into a direct measurement of spatiotemporal entomological risk (with advantages over the simpler “vector index” (21), if data on spatial or temporal variability in feeding patterns or vector competence are available) and provides a method for combining data from diverse mosquito species into a single continuous index. One difficulty with using this measure across different locations is that local data for mosquito vector competence or feeding patterns are rarely available, and both factors may be spatially variable (24–26), partly because of mosquito genetic factors (27, 28).

An additional recent challenge in predicting the risk of human WNV infection is the substantial decrease in federal and local funding for WNV surveillance. This raises the possibility that further cuts will result in little or no funds for WNV testing. Here, we attempt to provide validation and guidance for the use of mosquito surveillance data by determining their predictive power, with and without current WNV prevalence data, in 15 counties in Colorado, where 3,726 human WNV cases were reported between 2003 and 2007.

MATERIALS AND METHODS

In 15 counties in Colorado (Web Figure 1 available at <http://aje.oxfordjournals.org/>), carbon dioxide-baited CDC light traps and CDC gravid traps baited with a hay infusion were run for 1 night per week between May and October from 2003 to 2007 with some variability in starting and ending dates between counties; however, mosquito trapping took place in nearly all counties between June and September in the years when surveillance was active. Mosquitoes caught in traps were killed and identified by species, and *Culex* mosquitoes (including *Culex tarsalis*, *Culex pipiens*, *Culex restuans*,

and *Culex erythrothorax*) were pooled in groups of up to 50 and tested for WNV by reverse transcriptase–polymerase chain reaction (29). Mosquitoes from the 2 trap types were combined to reduce testing costs, and this practice may have added variability that reduced correlations and predictive power.

Records of human WNV infections were obtained from the Colorado Department of Public Health and Environment, and cases were aggregated to the county level to protect individual privacy. In our analyses, we included all types of WNV infection for a total of 3,726 reported cases. Most WNV cases were classified as fever (78.3%), followed by meningitis (11.8%), encephalitis (6.69%), asymptomatic status (1.58%), meningo-encephalitis (1.14%), and unknown status (0.49%). Although West Nile neuroinvasive disease is the most severe outcome, fever caused by WNV can last for several months and have lasting sequelae (12) and, thus, can also present a significant health burden at the individual and community levels.

We determined the utility of the mosquito surveillance data by correlating a vector risk index with the number of human WNV cases. Although a risk index that incorporates data on abundance, infection, feeding patterns, and vector competence (30) would be more accurate in predicting human cases, only limited data exist on spatial and temporal variations in mosquito feeding patterns in Colorado and none exist for vector competence (a limitation that many local and state agencies face). Surprisingly, our analyses suggest that human feeding frequency and partial vector competence did not differ significantly among *Culex* species (Web Figure 2) (31). As a result, we used a simplified vector index that was the sum across n mosquito species of the abundance, A_i (mosquitoes per trap-night), multiplied by the WNV infection prevalence, P_i , for that species, i , for that time step (e.g., weekly) as follows:

$$\text{Vector index} = \sum_{i=1}^n A_i P_i$$

This generalizes the “vector index” that has been used for single mosquito species in previous studies (21, 32) and by some counties. This simpler index will be almost as accurate as the fuller risk index described above as long as substantial differences in feeding and competence do not exist between the species being combined. However, if both *Aedes* and *Culex* mosquitoes are found to be infected, combining them into a risk index should take into account the substantial differences in feeding and vector competence (20).

We performed several analyses to determine the predictive power and quantitative relationship between the vector index (equation 1) and human WNV infections (Web Appendix). These included several spatial and temporal aggregations of testing and abundance data and analyses to determine the impact of not having WNV testing data from the current year, as would be the case under extreme resource limitations.

RESULTS

There were 3,827 cases of WNV reported in the 15 counties we studied in Colorado from 2003 to 2007, with the timing of seasonal peaks varying slightly from year to year and with substantial variability among counties (Web Figure 3, Web Table 1). The total number of cases in a county was not

correlated with the population size ($r=0.20$; $P=0.48$) or density ($r=-0.02$; $P=0.95$). Over the 5 years, 265,391 *Culex* mosquitoes were trapped and tested for WNV in the 15 study counties, with the majority being *C. tarsalis* (75%) and *C. pipiens* (20%); many fewer *C. erythrothorax* (2.7%), *C. restuans* (1%), and *Culex* spp. (1%) were trapped (Web Figure 4, Web Table 2). There were 922 WNV-positive pools, and most were *C. tarsalis* (77%) and *C. pipiens* (21%), with only a few positive pools from the other 2 species (Web Table 2).

Within a county, mosquito surveillance data integrated into the vector index described above were correlated with temporal variation in human WNV cases in most counties (Web Table 3, Figures 1 and 2, Web Figure 5), and correlations were higher in counties where large numbers of human cases occurred and where large numbers of mosquitoes were trapped (e.g., Larimer, Boulder, and Weld counties) (Figure 2). Overall, the 2 best methods to estimate prevalence and explain numbers of human WNV infections were the statewide weekly and local 2-week prevalence estimates (Web Appendix). Both of these measures were significantly correlated with human WNV cases 1–3 weeks in advance of the date of onset of illness (which occurs 3–14 days after infection) in 11–14 of the 15 counties, which accounted for 88%–98% of reported cases in the study area (Web Table 3). Both measures also had strong explanatory power up to 3 weeks in advance of the date of onset (pseudo- $R^2=0.45$ – 0.53). The vector index that used the statewide weekly prevalence estimate was frequently a better predictor than the local prevalence measures because it enabled risk estimates early and late in the season when few (or no) mosquitoes were trapped and prevalence could not be estimated locally.

In contrast, when current-year prevalence data were not used (simulating a situation in which funds are lacking to test current-year samples), mosquito surveillance data had very little explanatory power (Web Table 3). This poor fit resulted from substantial year-to-year variations in mosquito prevalence and the timing of peak prevalence that were not captured in the vector index without current-year data.

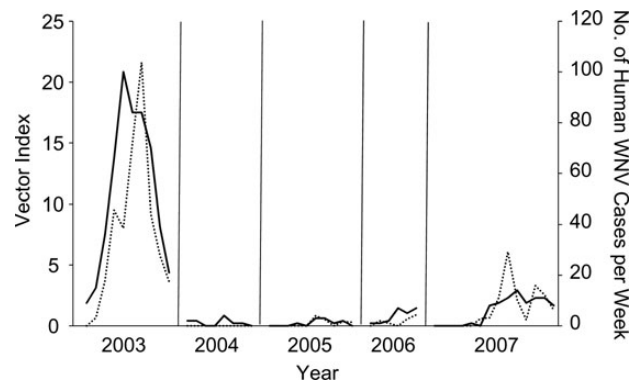


Figure 1. Vector index (dotted line) and human West Nile virus (WNV) cases (solid line) each week in Weld County, Colorado, 2003–2007. The vector index is the number of WNV-infected mosquitoes per trap-night and, in this figure, is calculated by using a local 2-week estimate of prevalence. Axes are square root–transformed to equalize leverage, linearize the relationship, and facilitate presentation.

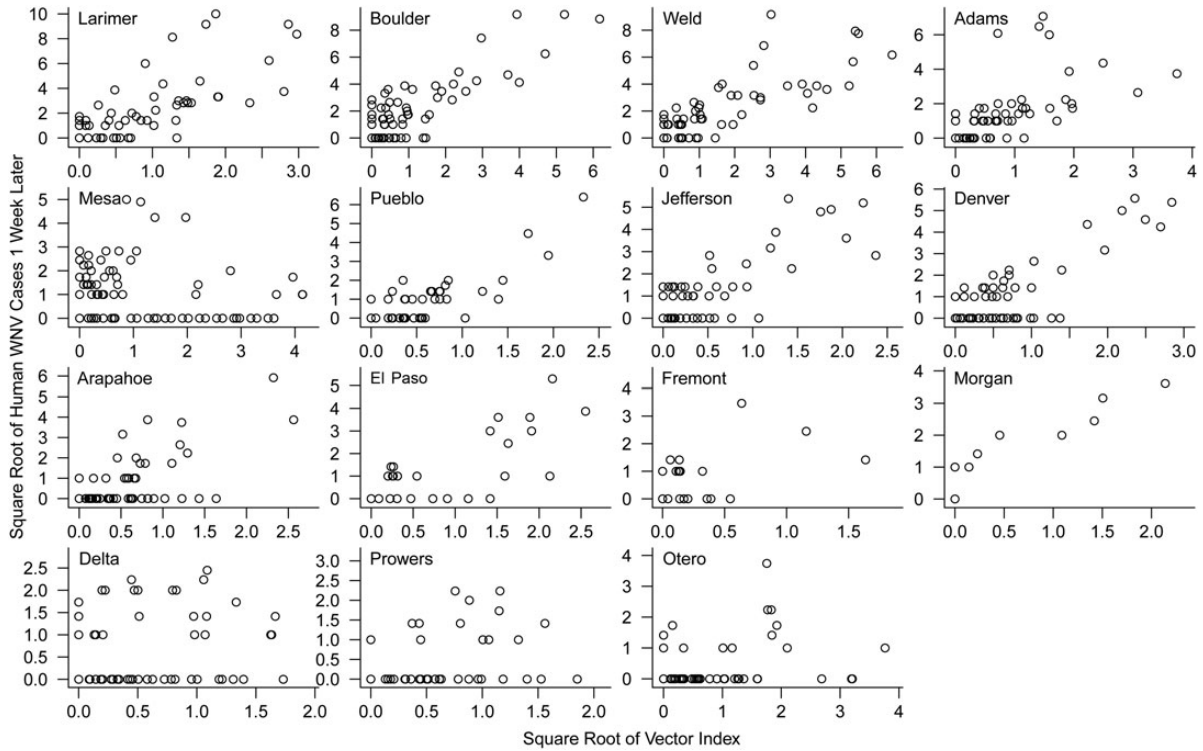


Figure 2. Human West Nile virus (WNV) cases 1 week later plotted against the vector index (with a local 2-week estimate of prevalence) for 15 counties in Colorado, 2003–2007, ordered from left to right and top to bottom in order of total number of cases over the 5 years. Each point is 1 week, and axes are square root–transformed to equalize leverage, linearize the relationship, and facilitate presentation. (See Web Table 3 for quantification of these correlations and for the strength of correlation 2 and 3 weeks in advance.)

Statewide spatiotemporal variation in the number of WNV cases was also strongly correlated with risk indices based on mosquito surveillance data (Figures 3 and 4A). Similar to

the local analyses, the most correlated vector indices used the statewide weekly method ($\text{pseudo-}R^2 = 0.53$) followed by the local 2-week method ($\text{pseudo-}R^2 = 0.44$). Much of the noise

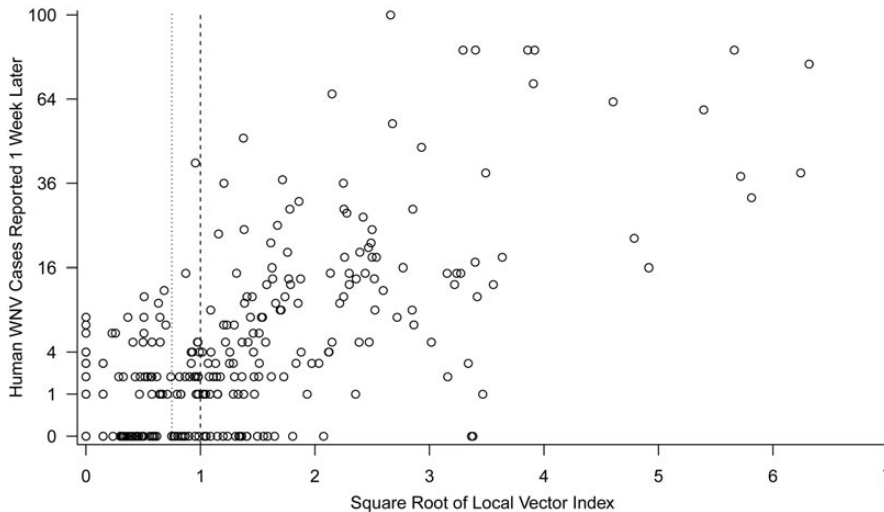


Figure 3. Number of West Nile virus (WNV) cases reported (on a square root–transformed scale) versus the square root of a 3-week moving window of the vector index 1 week prior, calculated by using a local 2-week prevalence estimate. Each point represents a county-week.

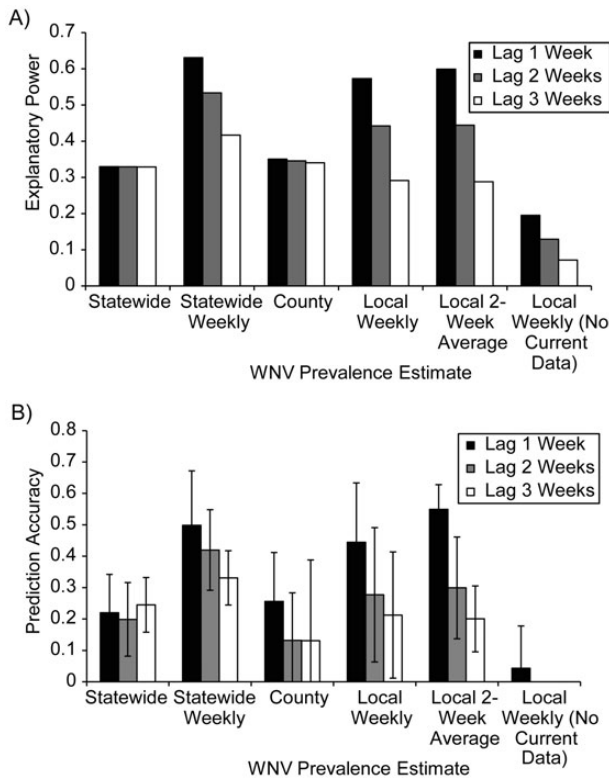


Figure 4. Statewide analysis of human West Nile virus (WNV) cases versus vector indices. A) Explanatory power (pseudo- R^2) of vector indices using various methods for estimating prevalence at 3 different time lags in the statewide analysis. B) Prediction accuracy ($1 - (\text{sum of squared residuals}) / (\text{total sum of squares})$) of risk indices when excluding a 10% random subset of the data and then using the fitted model to predict the number of WNV human cases by using the excluded mosquito surveillance data. Mean prediction accuracy within 1 standard deviation from 100 random subsets is shown. For “Local Weekly (No Current Data),” the mean prediction accuracy for lags of 2 and 3 weeks was 0.

(error) in these relationships resulted from substantial week-to-week variation in the risk index, such that both correlations could be substantially improved by using a 2- or 3-week moving window of the vector index. If moving windows of 2–3 weeks were used, then the vector index that used a local 2-week estimate of prevalence (pseudo- $R^2 = 0.89$) (Figure 3) was far superior to using the statewide weekly prevalence estimate (pseudo- $R^2 = 0.68$). Although a 3-week moving window (centered on the current week) results in a loss of 1 week in advance notice, it was still superior to using a 2-week moving window (the current and past week’s) for explaining the same week’s human cases (e.g., for the local 2-week vector index, using a 3-week window 2 weeks in advance was superior to using a 2-week window 1 week in advance (pseudo- $R^2 = 0.85$ and 0.78 , respectively). Again, explanatory power was poor for a vector index that did not include the average current-year prevalence data (pseudo- $R^2 = 0.13$) (Figure 4A).

We also assessed the power of the risk indices to predict future temporal and spatiotemporal variations in human

WNV cases by removing 10% of the data, fitting the model, and using this fitted model and the excluded mosquito surveillance data to predict the excluded human WNV case data (Figure 4B, Web Figure 5). The predictive accuracy was highest 1 week in advance for the statewide weekly and local 2-week prevalence methods ($R^2_{\text{pred}} > 50\%$) but was still substantial up to 3 weeks in advance of symptom onset.

The relationship between the local vector index and human cases could be used to set escalating thresholds for management actions. For example, the statewide data in Colorado (Figure 3) show that at least 1 human case frequently occurred in the week after the square root of the vector index was greater than 0.75 (calculated by using a 3-week moving window and a 2-week local prevalence estimate; positive predictive value = 86%) (Figure 3). Few cases occurred when this risk index was below 0.75 (negative predictive value (0 cases) = 57%; negative predictive value (≤ 1 cases) = 75%). Epidemic conditions, such as 4 or more cases per week, occurred frequently when the square root of the vector index was greater than 1 (positive predictive value = 69%) (Figure 3) and almost never at vector index values less than 1 (negative predictive value = 90%).

DISCUSSION

Mosquito control programs currently operate with limited resources, and recent WNV epidemics in 2012 are a strong reminder that a lack of effective control can result in substantial human illness. A key challenge has been identifying a surveillance method that would signal impending human infection and be reliable enough to be used to justify costly or unpopular control efforts, such as aerial spraying or curfews.

Our results show that standardized mosquito surveillance provides strong predictive power to signal human WNV infection up to several weeks in advance and is a valuable tool for public health officials. Mosquito data were correlated with both temporal variation in human WNV cases at local scales (over time within a county) and spatiotemporal variation at larger scales (e.g., statewide). These analyses build on previous studies of temporal correlations between vector indices and the number of human cases in Maryland in 2004 (23) and spatial correlations of vector indices and human WNV cases in Colorado in 2007 (21). They are a substantial advance over a recent study (18) that focused on predicting the presence or absence of human WNV cases but not their number. Our analyses are fully spatiotemporal and are based on a much larger data set (covering 5 years, a much larger area, and >3,800 reported human WNV cases) than those used in all previous studies and therefore provide greater confidence and justification for using the vector index as a guide for control actions (20, 21, 32). Given the 2012 WNV epidemics, our most promising finding is that high values of the risk index (Figure 3) have extremely high negative predictive value and strong positive predictive value. This suggests that waiting to take costly actions to control WNV when vector risk indices are low will nearly always be prudent, and taking costly actions to control WNV when risk indices are high is strongly warranted. It is worth noting that determining “high” and “low” values of vector indices requires analyses of data from the area where control actions are being considered.

Our results also highlight the lost value of letting surveillance efforts lapse. Predictive power was greatly reduced without WNV prevalence data from the current year (Figure 4) and would be far worse without mosquito abundance data from the current year (results not shown). The decay in predictive power that we observed with increasing advance warning (Web Table 3, Figure 4, Web Figure 5) also highlights the need to minimize delays in mosquito processing and testing and in implementing interventions.

Our results also provide guidance on the most effective way to estimate the prevalence of vector infection by using mosquito trapping and testing data, which can be challenging for arboviral surveillance because of the very low prevalence frequently observed. If local trapping efforts yield too few mosquitoes to accurately estimate prevalence, then the vector index should be calculated by using prevalence estimates for each mosquito species from regional or statewide trapping efforts for that week. In contrast, if trapping efforts over a 2-week period within a single county produce sufficient numbers of mosquito pools to adequately estimate WNV infection prevalence (e.g., at least 5 pools of 20 or more mosquitoes each), then a vector index using a 2-week estimate of prevalence can be used to predict human infection up to 3 weeks later. Further, our results suggest that using a “smoothed” or 3-week moving window of the vector index reduces the substantial week-to-week variability in mosquito abundance and prevalence and results in a more accurate risk index.

The strength of the associations between human cases and the vector index is heartening given recent findings suggesting that substantial variation in WNV incidence in humans exists within counties at the census-tract scale (33), and that differences in human behavior can sometimes be more important than differences in entomological risk (32). Our results suggest that, at least in Colorado, the link between WNV entomological risk and human risk is strong enough that predictive relationships are not obscured by unmeasured variation in these other factors. They also suggest that even a single set of mosquito traps placed in each county (as was the case for the data we analyzed) can provide useful information for allocating disease control efforts.

The epidemics of WNV in 2012 highlight an outstanding question that may be slightly more tractable given our results: What drives spatiotemporal variation in WNV infection in humans? The correlations presented here suggest that substantial insight can be gained from understanding the drivers of mosquito abundance and infection rates, although human behavior and immunity should not be ignored. Evidence suggests that both climate (34, 35) and urbanization (36–40) influence WNV transmission, but their relative importance and the quantitative relationships between these factors and WNV transmission are not yet known. Answering these broad-scale questions requires continued local surveillance efforts to track spatiotemporal variation in WNV transmission.

Our findings here, and those of others, indicate that active entomological surveillance provides a robust and valuable method to determine the risk of human infection with arboviruses like WNV. Resources to support these efforts and the trained personnel required to carry them out should not

be reduced by budget tightening. To do so in an era of environmental change and rapid international movement of vector-borne pathogens (41) could be perilous.

ACKNOWLEDGMENTS

Author affiliations: Department of Ecology and Evolutionary Biology, University of California, Santa Cruz, California (A. Marm Kilpatrick); and Colorado Department of Public Health and Environment, Denver, Colorado (W. John Pape).

Funding was provided by the National Science Foundation (grant EF-0914866), the National Institutes of Health (grant 1R01AI090159), and the National Institute of Allergy and Infectious Diseases (contract NO1-AI-25490).

We thank the numerous technicians who collected, organized, and tested the mosquito and human surveillance data and T. L. Bogich for technical assistance.

Conflict of interest: none declared.

REFERENCES

1. Kruse H, Kirkemo A, Handeland K. Wildlife as source of zoonotic infections. *Emerg Infect Dis.* 2004;10(12):2067–2072.
2. Bogich TL, Chunara R, Scales D, et al. Preventing pandemics via international development: a systems approach. *PLoS Med.* 2012;9(12):e1001354.
3. Centers for Disease Control and Prevention. West Nile virus: guidelines for surveillance, prevention, and control, 3rd revision. Fort Collins, CO: Centers for Disease Control and Prevention; 2003. (<http://www.cdc.gov/ncidod/dvbid/westnile/index.htm>). (Accessed December 8, 2004).
4. Ogden NH, St-Onge L, Barker IK, et al. Risk maps for range expansion of the Lyme disease vector, *Ixodes scapularis*, in Canada now and with climate change. *Int J Health Geogr.* 2008;7:24.
5. Randolph SE. The shifting landscape of tick-borne zoonoses: tick-borne encephalitis and Lyme borreliosis in Europe. *Philos Trans R Soc Lond B Biol Sci.* 2001;356(1411):1045–1056.
6. Süss J, Béziat P, Schrader C. Viral zoonosis from the viewpoint of their epidemiological surveillance: tick-borne encephalitis as a model. *Arch Virol Suppl.* 1997;13:229–243.
7. Petersen LR, Hayes EB. Westward ho? The spread of West Nile virus. *N Engl J Med.* 2004;351(22):2257–2259.
8. Hayes EB, Komar N, Nasci RS, et al. Epidemiology and transmission dynamics of West Nile virus disease. *Emerg Infect Dis.* 2005;11(8):1167–1173.
9. Kilpatrick AM. Globalization, land use, and the invasion of West Nile virus. *Science.* 2011;334(6054):323–327.
10. Centers for Disease Control and Prevention. West Nile virus. Fort Collins, CO: Centers for Disease Control and Prevention; 2013. (<http://www.cdc.gov/ncidod/dvbid/westnile/index.htm>). (Accessed January 16, 2013).
11. Health Canada. West Nile virus. Ottawa, Canada: Health Canada; 2012. (<http://www.phac-aspc.gc.ca/wnv-vwn/>). (Accessed January 16, 2013).
12. Kramer LD, Li J, Shi PY. West Nile virus. *Lancet Neurol.* 2007;6(2):171–181.

13. Marfin AA, Petersen LR, Eidson M, et al. Widespread West Nile virus activity, eastern United States, 2000. *Emerg Infect Dis.* 2001;7(4):730–735.
14. Guptill SC, Julian KG, Campbell GL, et al. Early-season avian deaths from West Nile virus as warnings of human infection. *Emerg Infect Dis.* 2003;9(4):483–484.
15. Theophilides CN, Ahearn SC, Grady S, et al. Identifying West Nile virus risk areas: the dynamic continuous-area space-time system. *Am J Epidemiol.* 2003;157(9):843–854.
16. LaDeau SL, Kilpatrick AM, Marra PP. West Nile virus emergence and large-scale declines of North American bird populations. *Nature.* 2007;447(7145):710–713.
17. Hochachka WM, Dhondt AA, McGowan KJ, et al. Impact of West Nile virus on American crows in the northeastern United States, and its relevance to existing monitoring programs. *EcoHealth.* 2004;1(1):60–68.
18. Kwan JL, Park BK, Carpenter TE, et al. Comparison of enzootic risk measures for predicting West Nile disease, Los Angeles, California, USA, 2004–2010. *Emerg Infect Dis.* 2012;18(8):1298–1306.
19. Patnaik JL, Juliusson L, Vogt RL. Environmental predictors of human West Nile virus infections, Colorado. *Emerg Infect Dis.* 2007;13(11):1788–1790.
20. Kilpatrick AM, Kramer LD, Campbell S, et al. West Nile virus risk assessment and the bridge vector paradigm. *Emerg Infect Dis.* 2005;11(3):425–429.
21. Bolling BG, Barker CM, Moore CG, et al. Seasonal patterns for entomological measures of risk for exposure to *Culex* vectors and West Nile virus in relation to human disease cases in northeastern Colorado. *J Med Entomol.* 2009;46(6):1519–1531.
22. Winters AM, Bolling BG, Beaty BJ, et al. Combining mosquito vector and human disease data for improved assessment of spatial West Nile virus disease risk. *Am J Trop Med Hyg.* 2008;78(4):654–665.
23. Kilpatrick AM, Kramer LD, Jones MJ, et al. West Nile virus epidemics in North America are driven by shifts in mosquito feeding behavior. *PLoS Biol.* 2006;4(4):606–610.
24. Kilpatrick AM, Daszak P, Jones MJ, et al. Host heterogeneity dominates West Nile virus transmission. *Proc Biol Sci.* 2006;273(1599):2327–2333.
25. Molaei G, Andreadis T, Armstrong P, et al. Host feeding patterns of *Culex* mosquitoes and West Nile virus transmission, northeastern United States. *Emerg Infect Dis.* 2006;12(3):468–474.
26. Hamer GL, Kitron UD, Brawn JD, et al. *Culex pipiens* (Diptera: Culicidae): a bridge vector of West Nile virus to humans. *J Med Entomol.* 2008;45(1):125–128.
27. Huang SM, Molaei G, Andreadis TG. Genetic insights into the population structure of *Culex pipiens* (Diptera: Culicidae) in the northeastern United States by using microsatellite analysis. *Am J Trop Med Hyg.* 2008;79(4):518–527.
28. Kilpatrick AM, Kramer LD, Jones MJ, et al. Genetic influences on mosquito feeding behavior and the emergence of zoonotic pathogens. *Am J Trop Med Hyg.* 2007;77(4):667–671.
29. Shi PY, Kauffman EB, Ren P, et al. High-throughput detection of West Nile virus RNA. *J Clin Microbiol.* 2001;39(4):1264–1271.
30. Kent R, Juliusson L, Weissmann M, et al. Seasonal blood feeding behavior of *Culex tarsalis* (Diptera: Culicidae) in Weld County, Colorado, 2007. *J Med Entomol.* 2009;46(2):380–390.
31. Farajollahi A, Fonseca DM, Kramer LD, et al. Bird-biting mosquitoes and human disease: a review of the role of *Culex pipiens* complex mosquitoes in epidemiology. *Infect Genet Evol.* 2011;11:1577–1585.
32. Gujral IB, Zielinski-Gutierrez EC, LeBailly A, et al. Behavioral risks for West Nile virus disease, northern Colorado, 2003. *Emerg Infect Dis.* 2007;13(3):419–425.
33. Winters AM, Eisen RJ, Delorey MJ, et al. Spatial risk assessments based on vector-borne disease epidemiologic data: importance of scale for West Nile virus disease in Colorado. *Am J Trop Med Hyg.* 2010;82(5):945–953.
34. IPCC Working Group I. *Climate Change 2007—The Physical Science Basis; Contribution of Working Group I to the Fourth Assessment Report of the IPCC.* Geneva, Switzerland: Cambridge University Press; 2007.
35. Kilpatrick AM, Meola MA, Moudy RM, et al. Temperature, viral genetics, and the transmission of West Nile virus by *Culex pipiens* mosquitoes. *PLoS Pathog.* 2008;4(6):e1000092.
36. Gomez A, Kilpatrick AM, Kramer LD, et al. Land use and West Nile virus seroprevalence in wild mammals. *Emerg Infect Dis.* 2008;14(6):962–965.
37. Andreadis TG, Anderson JF, Vossbrinck CR, et al. Epidemiology of West Nile virus in Connecticut: a five-year analysis of mosquito data 1999–2003. *Vector Borne Zoonotic Dis.* 2004;4(4):360–378.
38. Brown HE, Childs JE, Diuk-Wasser MA, et al. Ecological factors associated with West Nile virus transmission, northeastern United States. *Emerg Infect Dis.* 2008;14(10):1539–1545.
39. Bradley CA, Gibbs SEJ, Altizer S. Urban land use predicts West Nile Virus exposure in songbirds. *Ecol Appl.* 2008;18(5):1083–1092.
40. Pecoraro HL, Day HL, Reineke R, et al. Climatic and landscape correlates for potential West Nile virus mosquito vectors in the Seattle region. *J Vector Ecol.* 2007;32(1):22–28.
41. Kilpatrick AM, Randolph SE. Drivers, dynamics, and control of emerging vector-borne zoonotic diseases. *Lancet.* 2012;380(9857):1946–1955.

Web Appendix for

Predicting human West Nile virus infections with mosquito surveillance data

A. Marm Kilpatrick¹ and W. John Pape²

¹ University of California, Santa Cruz, CA, USA

² Colorado Dept. Public Health & Environment, Denver, CO, USA

Methods

Study Region

The fifteen counties in Colorado for which we had data for the analyses are shown in Web Figure 1, and the number of human cases in each county in each year is given in Web Table 1.

Mosquito feeding patterns

Data on the fraction of blood meals from mammals were available from Colorado (Weld and Larimer counties) for *Cx. tarsalis*, and *Cx. pipiens* (1), from California and Utah for *Cx. erythrothorax* (2-4), and from CT, MD/DC, NY/NJ, and TN for *Cx. pipiens* and *Cx. restuans* (5-8). Due to infrequent feedings on humans for *Cx. pipiens*, and *Cx. tarsalis* in the Colorado data (making it difficult to precisely estimate the fraction of blood meals from humans, F_h), we examined additional data on the fraction of mammalian blood meals that came from humans for these mosquito species from studies in California, Texas, and Utah for *Cx. tarsalis* (4, 9-13) and Kansas, California, Connecticut, New York, New Jersey, Tennessee, Maryland and Washington DC for *Cx. pipiens* (2, 5-8, 14, 15). Surprisingly, there was no significant difference in fraction of blood meals from humans or partial vector competence (Web Figure 2).

Estimating prevalence

One challenge in using the Vector Index is that the small number of mosquito traps used by most health departments at each site make it difficult to collect enough mosquitoes in a single week to accurately estimate WNV infection prevalence for each species. In addition, one of our key aims was to determine the predictive power of mosquito surveillance data when testing data was entirely missing, as may be the case under reduced budget conditions. Thus, we considered six ways of estimating prevalence P_i (estimated by maximum likelihood (16)) for each mosquito species i that we believe span potential strategies that might be used by local or state public health officials:

1. Statewide Static Prevalence: A statewide estimate, combining all the mosquitoes trapped over all weeks, all years, and across all counties. This essentially generates a risk index in which temporal variation is determined solely by mosquito abundance, with abundances of different mosquito species weighted by their statewide prevalence.
2. Statewide Weekly Prevalence: A statewide estimate from the week of trapping, resulting in different estimates for prevalence for each week of each year. This index ignores local (county) spatial variation in prevalence but explicitly includes temporal and between species differences in prevalence. It enables pooling of testing results across counties to decrease errors due to small numbers of mosquitoes tested, but, in doing so, obscures spatial variation in prevalence.
3. County Prevalence: An estimate for each county across all weeks. This index ignores temporal variation in prevalence, but incorporates among county variability.
4. Local Weekly Prevalence: An estimate from only the week of trapping in that county. Risk was calculated if the number of mosquitoes trapped was >40 ; otherwise the week was excluded from the analysis due to insufficient mosquitoes to estimate prevalence.

5. Local Two Week Prevalence: An estimate from the current week and previous week. As with Local Weekly Prevalence, risk was calculated if the number of mosquitoes trapped was >40; otherwise the week was excluded from the analysis.
6. Past Weekly Prevalence: An estimate from the calendar week of trapping in that county across all years except the “current” year. This estimate is a potential candidate for what counties might be forced to use in future years if no funds were available to test mosquitoes that year.

Statistical Analyses

We fit local data of the square root of counts of WNV cases vs. risk indices with generalized linear models with a quasi-poisson distribution and a square root link to equalize leverage and linearize relationships. We quantified the explanatory power of correlations with

$$\text{pseudo-}R^2 = 1 - \text{deviance}/\text{null deviance}$$

Pseudo- R^2 are approximations of the conventional R^2 but are more appropriate for non-Gaussian generalized linear models.

We fit statewide spatio-temporal numbers of WNV cases with a generalized linear mixed effect models with a poisson distribution and a square root link and with county as a random effect.

Results

Difference in key vectors

Although *Cx. tarsalis* was 3.65 times as abundant as *Cx. pipiens* overall, in two urban counties (Denver and Jefferson) *Cx. pipiens* was both more abundant and more frequently infected with WNV (Web Table 2), and in two other counties (Mesa, Pueblo), *Cx. pipiens* made

up more than 39% of the mosquitoes (with *Cx. tarsalis* making up most of the remainder). Finally, in one county (Delta), *Cx. erythrothorax* was nearly as abundant as *Cx. tarsalis*. Thus, while *Cx. tarsalis* is likely the most important WNV vector in Colorado for bird-to-bird, and bird-to-mammal transmission overall, *Cx. pipiens* may be more important in transmitting WNV to birds and mammals in some counties (Web Table 2), depending on the local feeding patterns of *Cx. pipiens* and *Cx. tarsalis* (see Discussion and Supplemental Material).

Comparison of prevalence methods

The correlation of mosquito surveillance data with the number of human WNV cases was much lower when using prevalence estimates that averaged across years within a county (Table 1: County) or across all years and counties (Table: Statewide), because substantial year to year variation was present both in calculated risk, and the number of human cases (Figure 1). Using a two-week running average prevalence measure outperformed using a single week estimate based on the average and case-weighted pseudo- R^2 values (Table 1). This is likely because using two weeks of trapped mosquitoes to estimate prevalence gave a more stable and accurate estimate of prevalence, while still capturing local variation in space and time.

Discussion

In collecting data for this analysis we were surprised by the variability observed in previous studies of mosquito feeding behavior, and discrepancies with conventional wisdom. Conventional wisdom suggested that *Cx. tarsalis* was a more mammalophilic vector and would thus feed more on humans than both *Cx. pipiens* and *Cx. restuans* which were thought to feed primarily on birds (17, 18). Instead, the average fraction of blood meals coming from humans across 24 studies, 550 – 13,600 blood meals/species, and 6-11 regions/species (Web Table 3), was highest for *Cx. restuans* (7.7% \pm 1 SE 3.5%; range 0-21%), followed by *Cx. pipiens* (4.4% \pm

1.4%; range 0-18%), *Cx. erythrothorax* ($2.7\% \pm 1.4\%$; range 0-5.4%), and *Cx. tarsalis* ($0.64\% \pm 2.7\%$; range 0-3.9%). What is clearly missing are studies determining the factors that influence mosquito feeding on humans and other mammals, or more generally, on all hosts broadly. Previous efforts have only considered temporal variation in feeding and have been either inconclusive (19), or have identified changes in the abundance of over-utilized hosts (American robins, *Turdus migratorius*) as predictors (20). Future studies should aim to identify causes of spatial variation in feeding patterns, especially with regard to humans. This would enable a more accurate estimating of the vector index that includes the likelihood of mosquitoes feeding on humans. This might help explain why we found no correlations of the number of human WNV cases with the population or population density within a county.

Conventional wisdom had also suggested that *Cx. tarsalis* was a more efficient vector in terms of vector competence than *Cx. pipiens* (17). However, recent results suggest this may not be the case when considering the standard measure of vector competence (the fraction of mosquitoes feeding on an infected blood meal that subsequently transmit) (21), and we found no difference in the part of vector competence relevant for estimating risk indices using WNV testing data (the fraction of infected mosquitoes that can transmit WNV) between the four species in this study (Web Figure 3).

Web Table 1. Numbers of reported human WNV cases and total population for 15 counties in Colorado from 2003 to 2007.

County	2003	2004	2005	2006	2007	Total	Population
Adams	238	15	4	12	32	301	363,857
Arapahoe	140	0	2	2	19	163	487,967
Boulder	457	14	5	76	98	650	291,288
Delta	10	27	1	34	6	78	27,834
Denver	173	3	5	5	29	215	554,636
El Paso	114	2	1	5	4	126	516,929
Fremont	77	4	4	1	15	101	46,145
Jefferson	160	8	6	8	34	216	527,056
Larimer	563	17	13	42	96	731	251,494
Mesa	20	127	10	38	37	232	116,255
Otero	28	0	1	6	10	45	20,311
Prowers	42	3	4	7	7	63	14,483
Pueblo	185	4	5	7	20	221	141,472
Weld	414	8	17	68	98	605	180,936
Morgan	64	0	0	2	14	80	28,109
All counties	2,685	232	78	313	519	3,827	3,568,772

Web Table 2. Number of mosquitoes trapped (WNV+ pools/pools tested) for each species for each county, across all weeks of 2004-2007. Prevalence (bottom row) for each species was estimated by maximum likelihood with bias correction.

Species	Cx.	Cx.	Cx.	Cx.	Cx.
County	<i>erythrothorax</i>	<i>Pipiens</i>	<i>restuans</i>	<i>spp.</i>	<i>Tarsalis</i>
Adams	0	687(1/42)	149(0/15)	0	471(0/26)
Alamosa	0	0	0	0	5208(1/111)
Arapahoe	0	203(1/29)	0	0	864(0/50)
Boulder	0	369(0/53)	0	0	2123(0/74)
Chaffee	0	8(0/1)	0	0	243(0/15)
Delta	2967(2/64)	108(0/10)	0	0	1302(4/45)
Denver	0	2814(0/101)	0	1(0/1)	575(1/39)
El Paso	0	26(0.2)	0	59(0/5)	145(0/9)
Fremont	0	0	0	0	196(0/14)
Jefferson	0	475(0/42)	0	0	68(0/17)
La Plata	0	0	0	0	45(0/5)
Larimer	0	455(0/47)	0	39(0/39)	2184(2/79)
Las Animas	0	137(2/13)	0	0	869(4/29)
Mesa	30(0/9)	207(7/50)	0	0	2717(23/101)
Otero	0	514(1/30)	0	0	3068(6/88)
Prowers	0	2(0/1)	0	0	1021(5/28)
Pueblo	0	457(1/35)	0	0	558(1/23)
Weld	0	389(1/42)	0	367(2/33)	3417(9/174)

All Counties	2997(2/73)	6851(14/498)	149(0/15)	466(2/51)	25074(56/927)
Prevalence	0.00067	0.0020	0	0.0035	0.0021

Web Table 3. Explanatory power of the vector index using six different methods for estimating prevalence for fifteen counties. The first three columns give the number of counties (out of 15 in the analysis) where the risk index was a significant predictor ($P < 0.05$) of the number of reported human cases, for all three lags (one, two, or three weeks in advance of the date of onset of illness), two of the three lags, or just one lag. The percentage in parentheses in the first column gives the percent of all human WNV cases that occurred in counties where the risk index was a significant predictor for all three lags. The next three columns give the average pseudo- R^2 for all fifteen counties for that lag (see Web Figure 3 for detailed results), and last column gives the weighted average pseudo- R^2 across all three lags for the Risk index for that prevalence estimate where the weights are the number of human cases in that county.

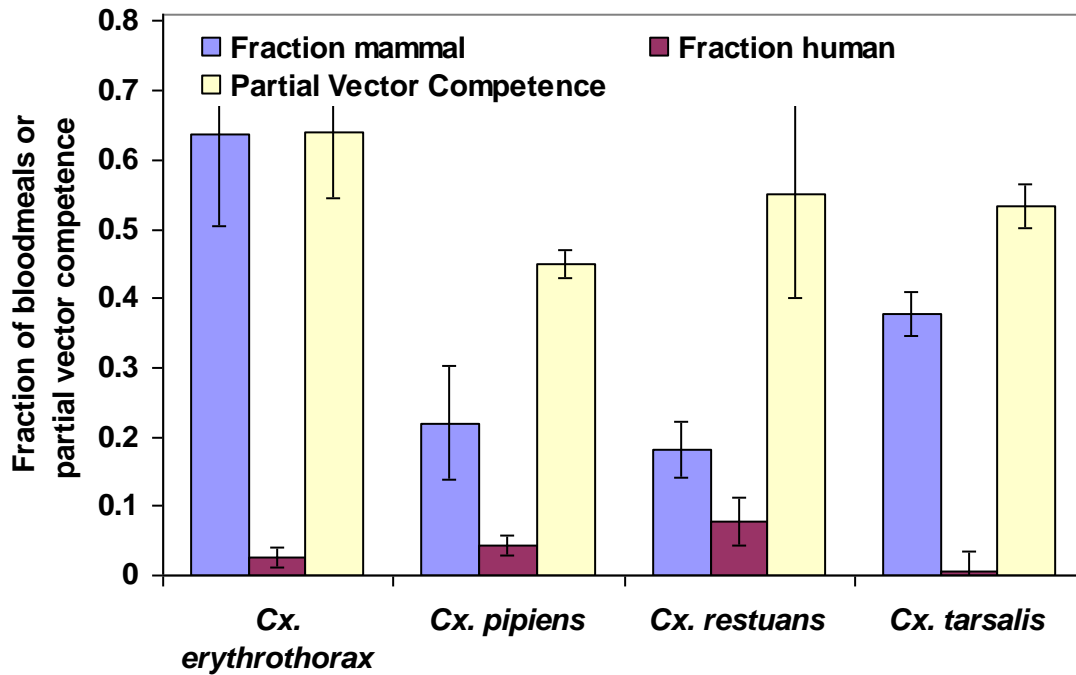
	Counties all 3 lags (% of human cases)	Counties 2 lags	Counties 1 lag	Average pseudo- R^2 , all counties			Case- weighted pseudo- R^2
				1 week lag	2 week lag	3 week lag	
Statewide	10 (90%)	2	0	0.30	0.29	0.32	0.33
Statewide							
Weekly	14(98%)	0	0	0.54	0.47	0.40	0.53
County	7(78%)	4	1	0.31	0.30	0.32	0.35
Local week	12(90%)	1	0	0.46	0.39	0.29	0.44
Local two-week	11(88%)	3	0	0.50	0.43	0.30	0.45
Local weekly (no	5(52%)	3	2	0.14	0.10	0.06	0.13

current year data)*

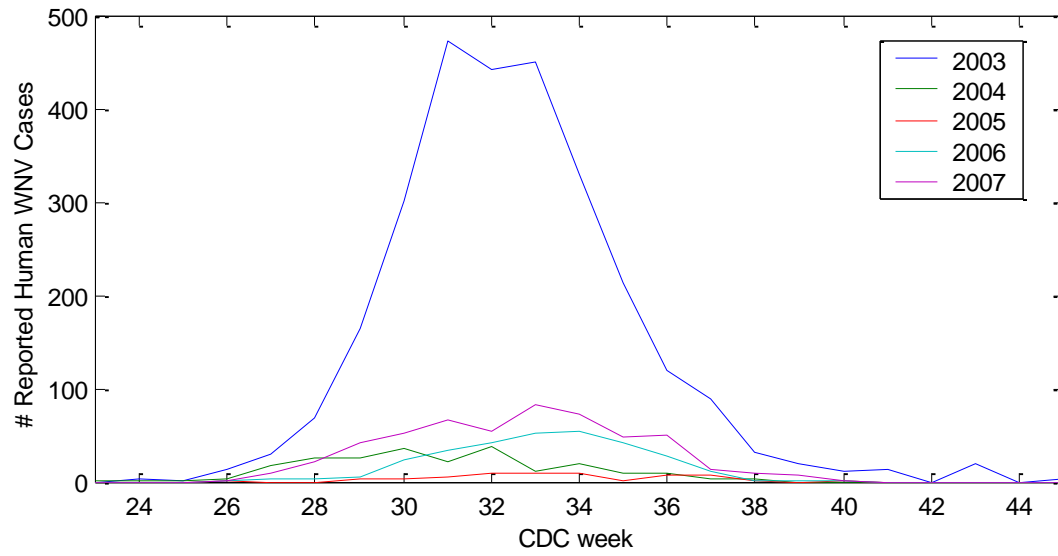
*One county only had sufficient mosquito surveillance data for analysis in one year (Morgan), so this risk measure could not be calculated and assessed. Thus total in 2nd column is out of 14 counties instead of 15.



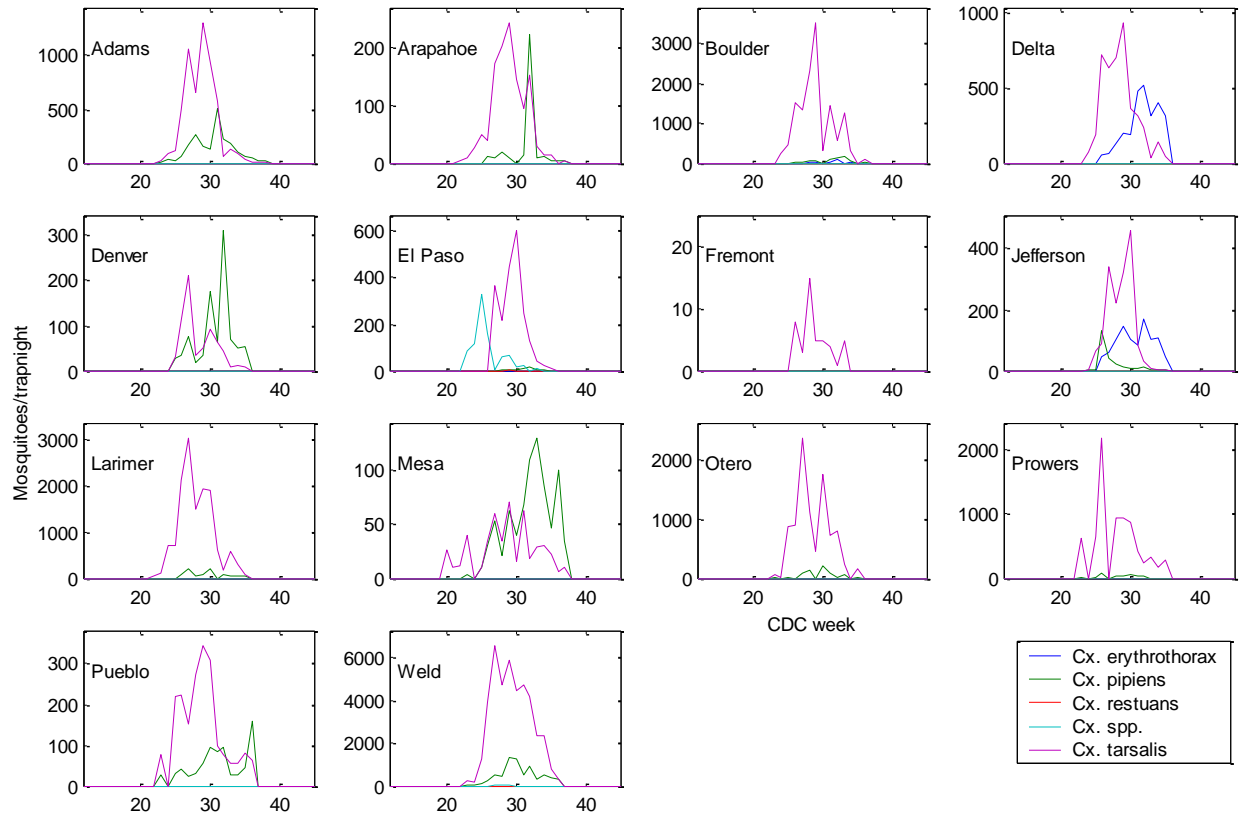
Web Figure 1. Map of Colorado counties. Shaded counties are those used in this study.



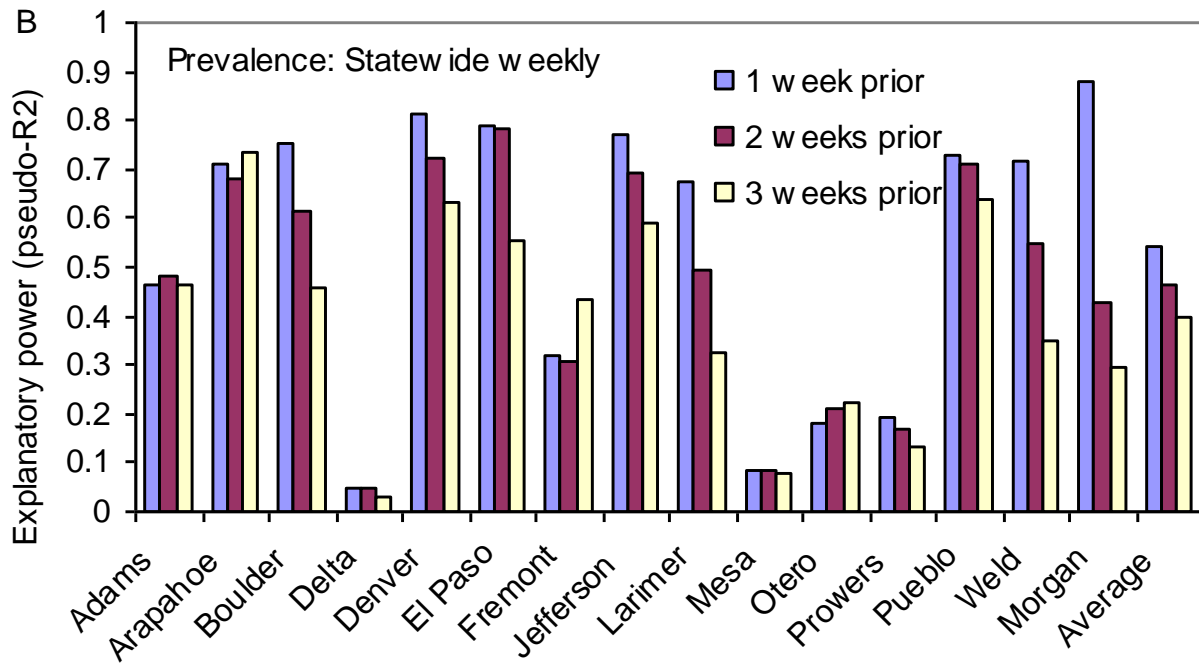
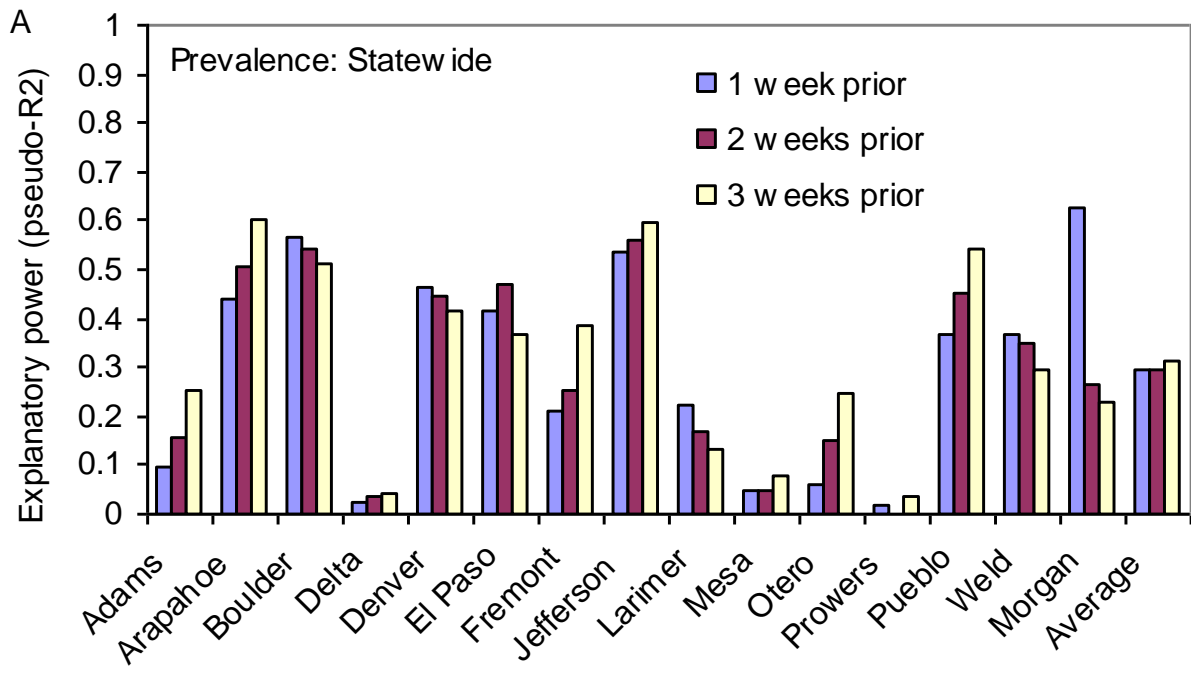
Web Figure 2. Feeding patterns and partial vector competence for four *Culex* mosquito species. Error bars show 1 standard error. The fraction of blood meals from mammals, F_m , differs significantly between mosquito species (ANOVA with arc-sin square-root transformed data: $F_{3,29}=4.47$; $P = 0.011$, but the fraction from humans, F_h , does not ($F_{3,25}=1.71$; $P = 0.19$). Partial vector competence, C_v , also did not differ significantly between species ($F_{3,11}$; $P = 0.93$).

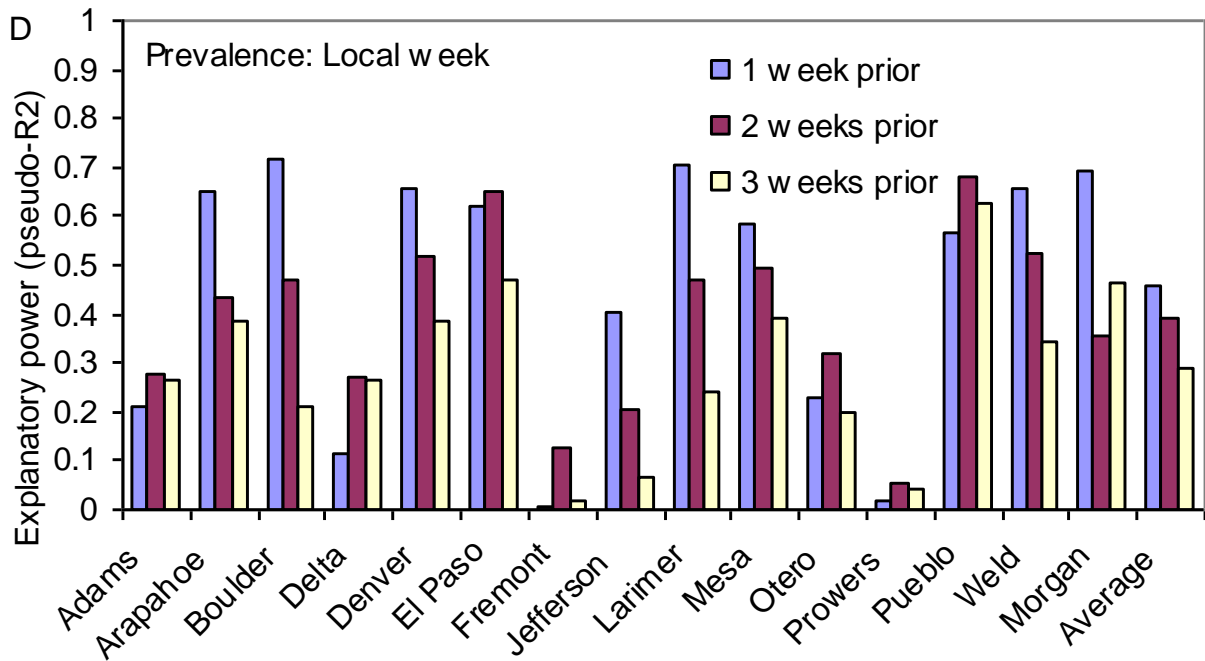
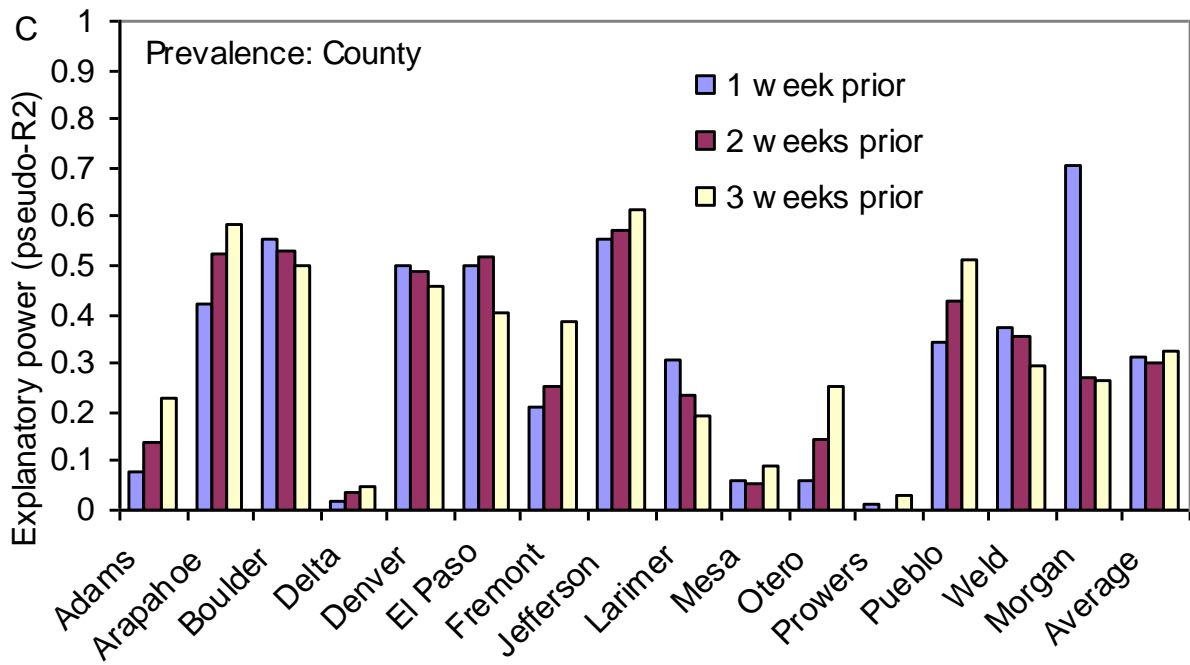


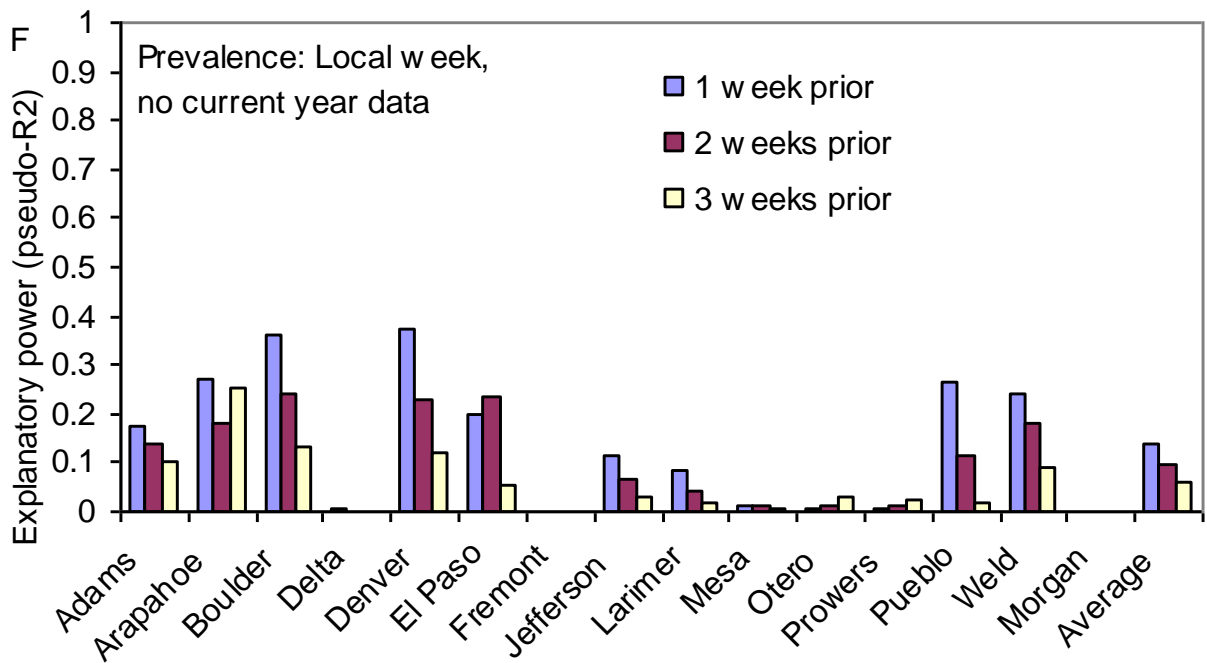
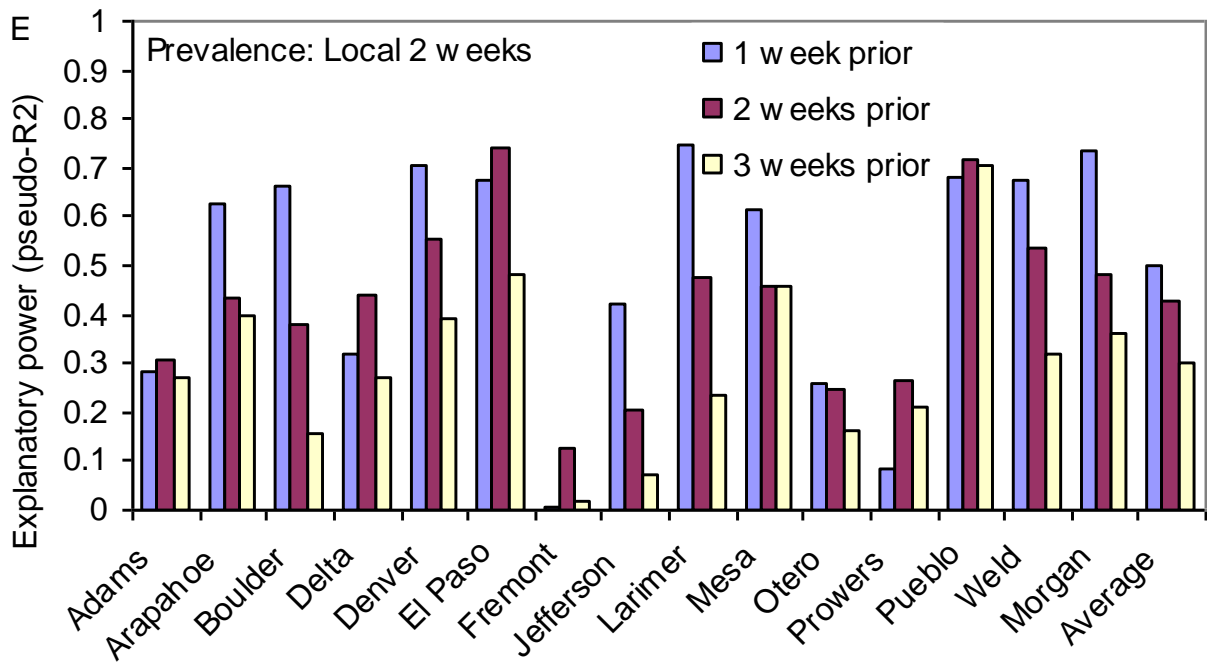
Web Figure 3. Number of human WNV cases reported each week in Colorado, from 2003 to 2007.



Web Figure 4. Mosquito abundance, by species, in fourteen counties in Colorado in 2007 vs. CDC week. Note different y-axis scales.







Web Figure 5. Predictive power (measured using the pseudo- R^2) of the square root of risk indices for predicting the square root of the number of human WNV cases using six different methods to estimate prevalence for fifteen counties in Colorado, 2003-2007. A) Statewide: prevalence estimate combines all mosquitoes of each species trapped over all years in all counties. It is essentially a sum of mosquito abundance weighted by the average prevalence for each species across the five years. Correlations using this risk index were significant ($P < 0.05$) for all counties for all lags, except: Fremont and Otero (which were significant for 2 and 3 week lags), Morgan (which was significant for 1 week lag only), and Delta and Prowers (which were non-significant for all lags). B) Statewide weekly: prevalence estimate combines all mosquitoes of each species trapped over all years in all counties for each week. This risk index was a significant predictor ($P < 0.05$) for all counties for all lags, except Delta, which was non-significant for all lags. C) County: prevalence was estimated using all mosquitoes trapped across all weeks in that county. This index was a significant predictor ($P < 0.05$) for eight of the fifteen counties for all lags, for one and two week lags in Arapahoe, for two and three week lags in Adams, Fremont and Otero, at one week lag in Morgan, and was non-significant for all lags in Prowers and Delta. D) Local week: prevalence was estimated using mosquitoes trapped in that county in that week, and risk is estimated if number of mosquitoes trapped is > 40 . This index was a significant predictor ($P < 0.05$) for twelve of the fifteen counties for all lags and for Jefferson for 1 and 2 week lags, but was non-significant for all lags for Fremont and Prowers. E) Local two weeks: prevalence was estimated using mosquitoes trapped in the current and previous week in that county. This index was a significant predictor ($P < 0.05$) for twelve of the fifteen counties for all 3 lags, for 1 and 2 week lags in Jefferson and Morgan, for 2 and 3

week lags in Prowers and was non-significant for all lags in Fremont. F) Local week, no current year data: prevalence was estimated using mosquitoes trapped in all years during the current week in the local county (similar to the statewide weekly index), but excludes testing results from mosquitoes trapped from the current year (simulating the situation where no funds are available to test mosquitoes locally). Risk was only a significant predictor ($P < 0.05$) for five of the fifteen counties for all 3 lags (Adams, Arapahoe, Boulder, Denver, Weld), for 1 and 2 week lags in El Paso, for a one week lag in Jefferson and Larimer, and was non-significant for all lags for five counties (Delta, Mesa, Otero, Prowers, Fremont). There was insufficient mosquito surveillance data from Morgan except in a single year (2003) so this prevalence method could not be used for this county.

Web Appendix References

1. Tempelis CH, Francy DB, Hayes RO, et al. Variations in feeding patterns of 7 culicine mosquitoes on vertebrate hosts in Weld and Larimer counties Colorado. *Am J Trop Med Hyg* 1967;16(1):111-119.
2. Gunstream SE, Chew RM, Hagstrum DW, et al. Feeding Patterns of 6 Species of Mosquitoes in Arid Southeastern California. *Mosquito News* 1971;31(1):99-101.
3. Tempelis CH. Estimation of vectorial capacity: mosquito host selection. *Bulletin of the Society for Vector Ecology* 1989;14(1):55-59.
4. Reisen WK, Milby MM, Presser SB, et al. Ecology of mosquitos and St. Louis Encephalitis Virus in the Los Angeles Basin of California, 1987-1990. *J Med Entomol* 1992;29(4):582-598.
5. Molaei G, Andreadis T, Armstrong P, et al. Host feeding patterns of *Culex* mosquitoes and West Nile virus transmission, northeastern United States. *Emerg Infect Dis* 2006;12(3):468-474.
6. Kilpatrick AM, Daszak P, Jones MJ, et al. Host heterogeneity dominates West Nile virus transmission. *Proceedings of the Royal Society B: Biological Sciences* 2006;273(1599):2327-2333.
7. Apperson CS, Hassan HK, Harrison BA, et al. Host feeding patterns of established and potential mosquito vectors of West Nile virus in the eastern United States. *Vector-Borne Zoonotic Dis* 2004;4(1):71-82.
8. Savage HM, Aggarwal D, Apperson CS, et al. Host choice and West Nile virus infection rates in blood fed mosquitoes, including members of the *Culex pipiens* complex, from

- Memphis and Shelby County, Tennessee 2002-2003. *Vector-Borne Zoonotic Diseases* 2007;7(3):365-386.
9. Hayes RO, Tempelis CH, Hess AD, et al. Mosquito host preference studies in Hale County, Texas. *Am J Trop Med Hyg* 1973;22(2):270-277.
 10. Andersen DM, Collett GC, Winget RN. Preliminary host preference studies of *Culex tarsalis* Coquillett and *Culiseta inornata* (Wiluston) in Utah. *Mosquito News* 1967;27(1):12-15.
 11. Tempelis CH, Reeves WC, Bellamy RE, et al. A 3-Year study of feeding habits of *Culex tarsalis* in Kern County California. *Am J Trop Med Hyg* 1965;14(1):170-177.
 12. Tempelis CH, Washino RK. Host-feeding patterns of *Culex tarsalis* in Sacramento Valley California with notes on other species. *J Med Entomol* 1967;4(3):315-318.
 13. Wekesa JW, Yuval B, Washino RK, et al. Blood feeding patterns of *Anopheles freeborni* and *Culex tarsalis* (Diptera: Culicidae): effects of habitat and host abundance. *Bull Entomol Res* 1997;87(6):633-641.
 14. Edman JD, Downe AER. Host-blood sources and multiple-feeding habits of mosquitoes in Kansas. *Mosq News* 1964;24(2):154-160.
 15. Apperson CS, Harrison BA, Unnasch TR, et al. Host-feeding habits of *Culex* and other mosquitoes (Diptera: Culicidae) in the Borough of Queens in New York City, with characters and techniques for identification of *Culex* mosquitoes. *J Med Entomol* 2002;39:777-785.
 16. Biggerstaff BJ. Confidence intervals for the difference of two proportions estimated from pooled samples. *Journal of Agricultural Biological and Environmental Statistics* 2008;13(4):478-496.
 17. Turell MJ, Dohm DJ, Sardelis MR, et al. An update on the potential of North American mosquitoes (Diptera : Culicidae) to transmit West Nile virus. *J Med Entomol* 2005;42(1):57-62.
 18. Turell MJ, Sardelis MR, O'Guinn ML, et al. Potential vectors of West Nile virus in North America. In: Mackenzie J, Barrett A, Deubel V, eds. *Japanese Encephalitis and West Nile Viruses Vol 267 Current Topics in Microbiology and Immunology*. Berlin: Springer-Verlag, 2002:241-252.
 19. Edman JD, Taylor DJ. *Culex Nigripalpus* - Seasonal Shift in Bird-Mammal Feeding Ratio in a Mosquito Vector of Human Encephalitis. *Science* 1968;161(3836):67-68.
 20. Kilpatrick AM, Kramer LD, Jones MJ, et al. West Nile virus epidemics in North America are driven by shifts in mosquito feeding behavior. *PLoS Biology* 2006;4(4):606-610.
 21. Reisen WK, Barker CM, Fang Y, et al. Does variation in *Culex* (Diptera: Culicidae) vector competence enable outbreaks of West Nile virus in California? *J Med Entomol* 2008;45(6):1126-1138.