Identification of hyperendemic foci of horses with West Nile virus disease in Texas

Courtney A. Wittich, MS; Michael P. Ward, BVSc, MSc, MPVM, PhD; Geoffrey T. Fosgate, DVM, PhD; Raghavan Srinivasan, PhD

Objective—To determine whether West Nile virus (WNV) disease hyperendemic foci (hot spots) exist within the horse population in Texas and, if detected, to identify the locations.


Procedures—Case data with spatial information from WNV epidemics occurring in 2002 (1,377 horses), 2003 (396 horses), and 2004 (134 horses) were analyzed by use of the spatial scan statistic (Poisson model) and kriging of empirical Bayes smoothed county attack rates to determine locations of horses with WNV disease in which affected horses were consistently (in each of the 3 study years) clustered (hyperendemic foci, or hot spots).

Results—2 WNV hot spots in Texas, an area in northwestern Texas and an area in eastern Texas, were identified with the scan statistic. Risk maps of the WNV epidemics were qualitatively consistent with the hot spots identified.


West Nile virus is a flavivirus endemic in many areas of Africa, western Asia, the Middle East, and, most recently, the United States. It is a vector-borne infection maintained in nature by a mosquito-bird cycle. In humans and horses, WNV infection is usually asymptomatic or characterized by a mild febrile illness. However, fatal meningoencephalitis or encephalitis does occur. For horses, a fatality rate of ≥30% has been reported.

Since its introduction to North America (New York City) in August 1999, the geographic range of WNV has increased substantially in a south and west direction across the United States. The number of reported horses with WNV disease dramatically increased in 2002, when 9,144 horses with WNV disease from 38 states were reported to the CDC.

Information from mosquito surveillance is important for designing WNV-disease prevention programs. Although the vectors of WNV that maintain the transmission cycle within wild bird reservoirs have not definitively been identified, agreement exists that the Culex genus is important for transmission to humans and horses. The occurrence of WNV disease in humans has been explained by changes in enzootic (bird-to-bird) and bridge (bird-to-human) vector feeding preferences, specifically those involving Culex pipiens in the north-east and north-central United States and Culex tarsalis in Colorado and California. It appears that C pipiens and Culex salinarius might be important vectors of WNV in horse populations in parts of the United States, although bridge vector species that transmit WNV from wild bird populations to horses remain unclear.

Environmental factors are likely to be important determinants of where WNV infections occur. In general, WNV transmission is likely to occur in areas that have abundant water (eg, near lakes and rivers) and during periods of higher temperatures (summer to early fall). However, there might be great variability in local environmental factors that determine the occurrence of WNV disease.

The link between the environment and WNV transmission suggests that the occurrence of WNV disease in horse populations should be highly clustered. Previous research has used geographic information-science technologies to investigate the spatial distribution of WNV-positive mosquito pools, locations of dead crows used within surveillance systems to detect the presence of

**Abbreviations**

<table>
<thead>
<tr>
<th>WNV</th>
<th>West Nile virus</th>
</tr>
</thead>
<tbody>
<tr>
<td>RR</td>
<td>Relative risk</td>
</tr>
</tbody>
</table>

Received September 11, 2007.
Accepted November 7, 2007.

From the Department of Veterinary Integrative Biosciences, College of Veterinary Medicine and Biomedical Sciences, Texas A&M University; College Station, TX 77843-4458 (Wittich, Ward, Fosgate); the Department of Ecosystem Science & Management, College of Agriculture and Life Sciences, Texas A&M University, College Station, TX 77843-2135; and the Spatial Sciences Laboratory, Texas A&M University, College Station, TX 77843-2120 (Srinivasan).

This manuscript represents a portion of a thesis submitted by the senior author to the Veterinary Medicine & Biomedical Sciences Graduate School as partial fulfillment of the requirements for a Master of Science degree at Texas A&M University.

Dr. Ward was supported by a grant from the Texas Equine Research Account Advisory Committee.

Address correspondence to Dr. Ward.
WNV, and locations of humans with WNV disease.\textsuperscript{13-17} There have been previous attempts to identify hyperendemic foci (hot spots) of WNV disease in horse populations.\textsuperscript{3,6,20,21} However, these studies have focused on small-scale WNV epidemics occurring over short periods.

Spatial analysis of WNV disease has been shown to aid in the spatial prediction of risk of infection.\textsuperscript{18-21} If high-risk areas can be identified, then preventive measures, such as mosquito control, vaccination of susceptible horses, and education of owners, can be implemented to decrease the impact of this important disease. By use of reported WNV-disease data from 2002 to 2004, the aim of the study reported here was to determine whether hot spots of WNV disease exist within the horse population in Texas and, if detected, to identify the locations.

**Materials and Methods**

**Source data**—Reports of WNV disease occurring in horses in Texas during 2002, 2003, and 2004 compiled by the Texas Department of State Health Services were accessed. All affected horses had clinical signs suggestive of WNV encephalomyelitis (ataxia, abnormal gait, muscle fasciculations, recumbency, and depression) and were confirmed to have WNV disease by positive IgM antibody capture-ELISA (titer > 1:400) results. Clinical signs and a positive IgM antibody capture-ELISA result are considered sufficient criteria for confirmation of WNV disease.\textsuperscript{4} West Nile virus–specific IgM antibodies are detectable 6 to 10 days after infection and persist for 2 to 3 months.\textsuperscript{22,4} Vaccination is unlikely to produce false-positive IgM antibody capture-ELISA results.\textsuperscript{4} Data available for affected horses reported in 2002, 2003, and 2004 included information on disease onset, residential address, sex, and age. Latitude and longitude coordinates were reported for some of the affected horses. Case-report data were initially compiled, organized, and checked within a spreadsheet program.\textsuperscript{9}

**Data analysis**—Data were imported into a geographic information system\textsuperscript{1} for further formatting and analysis. A shape file of Texas counties\textsuperscript{8} was overlaid with a polyline file of highways and streets.\textsuperscript{4} For horses with WNV disease that had reported latitude and longitude coordinates, the locations were added to the map of Texas counties. For horses with WNV disease that did not have reported latitude and longitude coordinates but which had a reported street address (consisting of street name, street type, street number, and zip code), locations were geocoded by use of an address locator.\textsuperscript{7} For some affected horses, location information was qualitative (for example, distance and direction from the nearest town on a main road) but locations could be identified visually, and latitude and longitude could be estimated. Horses with WNV disease with reported addresses that consisted of post office boxes or rural route numbers and affected horses with missing street address information or nonspecific location directions were excluded from data analysis. Locations of horses with WNV disease were projected by use of the North American Datum of 1983.

The number of reported affected horses per Texas county was calculated by use of a spatial query: The WNV-disease attack rate for each county was calculated by dividing the number of affected horses reported from each county by the number of horses at risk (population) in each county\textsuperscript{23} and was estimated as attack rate/1,000 horses at risk.

The mean center of reported horses (the location that represents the mean x-coordinate value and the mean y-coordinate value of all affected horses in Texas) was calculated for each study year.\textsuperscript{6} Directional ellipses (a measure of whether a spatial distribution of points has a directional trend) were also calculated for each study year by use of an output size of 1 SD. The procedures were repeated, weighting mean centers and directional ellipses by the date of onset (Julian day) for each affected horse so that horses reported later in each outbreak influenced the calculated means and directional ellipses more. The spatial patterns of affected horses reported in each year of the study were characterized by use of the Moran autocorrelation statistic,\textsuperscript{6} weighted by the reported date of disease occurrence. The Moran spatial autocorrelation statistic is a measure of the degree to which a set of spatial features and their associated data values (date of disease onset) tend to occur together (positive spatial autocorrelation) or not (negative spatial autocorrelation). It is a measure of spatial structure with respect to some attribute (date of disease occurrence) of interest.

A directional statistic\textsuperscript{24} was used to determine whether a systematic directional spread of outbreaks occurred within Texas during the study period. A chain of infection was constructed by first sequencing the outbreaks by date of occurrence (the primary outbreak first, followed by the second outbreak, etc). A line was then drawn to connect the location of the first outbreak to the location of the second outbreak, repeating this until all outbreaks were connected. The chain of infection had 2 ends (the first and last outbreaks) and branches when outbreaks occurred on the same date. Various possible chains of infection can be specified. In this study, the time-connection matrix was specified as adjacent, in which each outbreak connected only to its temporal nearest neighboring outbreak. The test statistic is a vector, the direction of which is the mean direction of the links making up the chain of infection, and the magnitude of which is the angular variance of the links. When the links all point in the same direction, the angular variance is small; when they point in many directions, the angular variance is large. The significance of the test statistic was estimated by use of Monte Carlo simulation.\textsuperscript{7} Geographic data were transformed for this analysis by use of the Universal Transverse Mercator system, projection zone 14. The Universal Transverse Mercator system is a projected coordinate system that divides the world into 60 north and south zones, 6° wide.\textsuperscript{6}

To identify specific hot spots of reported WNV disease in the 3 study years, data were analyzed by use of the space-time scan statistic.\textsuperscript{8} The scan statistic uses a window of variable size to scan data for clusters of disease instances or rates of disease occurrence and uses a likelihood ratio statistic to test whether such a
cluster could be explained by chance. Data were edited, and the following text files were created separately for each year of the study: case files (horse identification number, date of onset, and number of affected horses), population files (identification number, year, and population per county), and coordinate files (identification number, latitude, and longitude). Study period was defined as having a start date of year (2002, 2003, or 2004), month (01), and day (01) and an end date of year (2002, 2003, or 2004), month (12), and day (31).

Data were scanned for time-space clusters by use of a Poisson (population-at-risk model) with a scanning window of ≤ 30 days in length and ≤ 100 km in area. The population at risk was assumed to be the estimated county horse population. Data were only scanned for clusters with no geographic overlap. Each cluster was described by a center (latitude and longitude coordinates) and a radius. Center and radius information was imported into a geographic information system, and overlay analysis was used to identify Texas counties included within clusters identified in all 3 study years. These counties are referred to as hot-spot counties.

To observe WNV-disease risk and to validate counties identified as hot-spot counties, kriging was performed on the county WNV-disease attack-rate data. Kriging is an interpolation technique in which the surrounding measured values are weighted to derive a predicted value for an unmeasured location. Weights are based on the distance between the measured point, prediction locations, and overall spatial arrangement among the measured points. Kriging is unique among the interpolation methods in that it provides an easy method for characterizing the variance, or precision, of predictions. Kriging is based on regionalized variable theory, which assumes that the spatial variation in data being modeled is homogeneous across the surface. That is, the same pattern of variation can be observed at all locations on the surface. However, use of disease rates or proportions calculated for areas in which the population at risk or sampled varies substantially can introduce bias in estimated disease risk. For example, the horse population at risk in Texas counties varies from < 50 to approximately 10,000. In these circumstances, smoothing disease rates or proportions prior to interpolation can produce more robust and valid risk maps.

A spatial weight set of Texas counties was created. Polygon adjacency was defined by use of a queen neighbor criterion (ie, areas that share the edge to the immediate left, right, up, and down as well as diagonal edges) and the 10 nearest neighboring counties. The spatial weight for each county was standardized by the number of nearest neighbors. Crude equine county WNV-disease attack rates were smoothed with an empirical Bayes smoothing algorithm by use of the defined spatial weights and the estimated population at risk in each county.

The centroid of each Texas county was identified. By use of Texas county centroids to define spatial location, the semivariance between and smoothed county attack rates for all possible pairs of counties for each study year was calculated, and a semivariogram was constructed with 10 lags of 25 km. For each study year, models (Gaussian, spherical, exponential, and power) were fit to the semivariogram, and range, sill, and nugget variables were estimated. Estimated variables were used to fit ordinary kriging models' year-specific smoothed attack-rate data sets to produce maps of interpolated WNV-disease attack rates.

Raster outputs for each study year were layered with shape files of hot-spot counties to assess qualitatively the validity of the hot-spot locations identified.

Results

During the period of 2002 to 2004, 2,583 horses with WNV disease were reported (1,698 [65.7%], 717 [27.8%], and 168 [6.5%]) horses in 2002, 2003, and 2004, respectively). Overall, 1,907 affected horses had geographic coordinates (decimal degrees) reported or derived (1,377, 396, and 134 horses in 2002, 2003, and 2004, respectively). For affected horses with geographic coordinates, date of disease onset was reported for 1,371, 389, and 74 horses in 2002, 2003, and 2004, respectively. Median age of affected horses included (8 years) and excluded (7 years) from analysis in this study was not significantly (P = 0.162) different. The number of male and female horses included (652 and 558, respectively) was not significantly (P = 0.114) different from the number (120 and 128, respectively) excluded from analysis.

Horses with WNV disease were reported throughout Texas in 2002 and 2003, whereas affected horses were reported more commonly from eastern Texas during 2004 (Figure 1). Based on mean centers (Figure 2), the distribution of horses with WNV disease shifted from central Texas in a southeast direction during the period of 2002 to 2004. Although all estimated directional ellipses overlapped, all ellipses had a northwest-by-southeast distribution in the 3 series. The spatial pattern that was observed for reported date of disease onset was similar to that observed for the analysis without weighting by reported date of disease onset. The Moran autocorrelation statistic for reported date of disease onset for 2002, 2003, and 2004 was 0.29 (P < 0.001), 0.26 (P < 0.001), and 0.16 (P = 0.016), respectively.

Directionality was significant (P < 0.001) for the 3 epidemics. The mean direction of the 2002, 2003,
and 2004 epidemics, with temporal case adjacency and angles taken as counterclockwise degrees from horizontal with east corresponding to 0° and north to 90°, were 227° (southwest to northeast), 313° (southeast to northwest), and 265° (south to north), respectively. The angular concentrations, representing the variance in the angles between connected affected horses, for the 3 epidemics were similar (0.189, 0.129, and 0.202) and suggested in general a consistent direction of spread.

Two clusters of high WNV-disease attack rates were identified by use of the scan statistic in each of the study years as follows: 2002) latitude, –101.7428° S; longitude, 34.3458° E; radius, 99.54 km (RR, 15.792; P = 0.001) and latency, –97.2983° S; longitude, 32.9312° E; radius, 89.9 km (RR, 5.628; P = 0.001); 2003) latitude, –102.166° S; longitude, 32.700° E; radius, 99.83 km (RR, 21.87; P = 0.001) and latency, –97.075° S; longitude, 31.333° E; radius, 89.15 km (RR, 7.932; P = 0.001); and 2004) latitude, –102.0487° S; longitude, 31.96° E; radius, 87.42 km (RR, 64.244; P = 0.001) and latitude, –95.2125° S; longitude, 30.096° E; radius, 60.49 km (RR, 4.318; P = 0.030). When mapped and overlaid, clusters identified in each study year formed 2 hot spots as follows: 6 counties (Cochran, Hockley, Lubbock, Terry, Lynn, and Garza) located in northwestern Texas and 4 counties (Freestone, Limestone, Leon, and Robertson) located in eastern Texas (Figure 3).

For each study year, a comparison of estimated

![Figure 2](image-url)  
![Figure 3](image-url)
crude and empirical Bayes smoothed county WNV-disease attack rates was determined (Table 1). Smoothing resulted in a reduction in mean county attack rates but an increase in median county attack rates in each year. The variability of county attack rates was reduced (30% to 44%) in each year. The shapes of the distributions of attack rates in 2002 and 2003 were not substantially affected by smoothing, but the 2004 distribution of smoothed attack rates was more skewed, and kurtosis was increased by smoothing.

Each year-specific semivariogram of smoothed county attack rates of WNV disease was best fit by a Gaussian model. Resulting maps of interpolated WNV-disease attack rate (Figure 4) revealed a large area of high attack rates in northwestern Texas in 2002 and in southeastern Texas in 2004. Other hot spots were apparent in central and eastern Texas in 2002 and in eastern Texas in 2003. The northwestern and eastern hot spots identified by the scan-statistic analysis were consistent with hot spots observed in interpolated maps of WNV-disease attack rates for 2002 and 2004, respectively. In 2003, these hot spots bordered areas of high-interpolated WNV-disease attack rates. During the 3-year period, the northwestern hot spot appeared to become weaker, whereas the eastern hot spot apparently became stronger, relative to the overall attack rates estimated for Texas in each study year.

**Discussion**

Two hot spots of WNV disease in horses in Texas

<table>
<thead>
<tr>
<th>Attack rates per year*</th>
<th>Minimum</th>
<th>Maximum</th>
<th>Mean ± SD</th>
<th>Median</th>
<th>Skewness</th>
<th>Kurtosis</th>
</tr>
</thead>
<tbody>
<tr>
<td>2002</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Crude</td>
<td>0</td>
<td>57.4</td>
<td>5.75 ± 9.39</td>
<td>2.50</td>
<td>2.97</td>
<td>9.79</td>
</tr>
<tr>
<td>Smoothed</td>
<td>0.05</td>
<td>36.6</td>
<td>4.94 ± 6.60</td>
<td>2.63</td>
<td>2.29</td>
<td>5.21</td>
</tr>
<tr>
<td>2003</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Crude</td>
<td>0</td>
<td>32.8</td>
<td>1.42 ± 3.87</td>
<td>0</td>
<td>5.94</td>
<td>40.3</td>
</tr>
<tr>
<td>Smoothed</td>
<td>0.02</td>
<td>21.7</td>
<td>1.22 ± 2.39</td>
<td>0.46</td>
<td>5.68</td>
<td>39.8</td>
</tr>
<tr>
<td>2004</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Crude</td>
<td>0</td>
<td>9.61</td>
<td>0.35 ± 1.07</td>
<td>0</td>
<td>4.93</td>
<td>29.8</td>
</tr>
<tr>
<td>Smoothed</td>
<td>0.02</td>
<td>6.88</td>
<td>0.31 ± 0.60</td>
<td>0.13</td>
<td>6.46</td>
<td>58.5</td>
</tr>
</tbody>
</table>

*Horses with WNV disease were reported from 200 (79%), 121 (48%), and 54 (21%) of 254 Texas counties in 2002, 2003, and 2004, respectively.
were identified with the scan statistic and were corroborated by disease risk maps. Although reports of WNV disease in horses have shifted away from northwestern Texas in 2003 and 2004, a hot spot location identified by the scan statistic persisted in this region throughout the study period. The 6 counties included in this hot spot contain an estimated horse population of only 3,130, or 522 horses/county, which is considerably less than the mean county population of horses in Texas (1,466 horses). Further study of this hot spot could reveal much about the epidemiologic characteristics of WNV in horse populations. Possible explanations for the persistence of this hot spot might include specific environmental habitats ideal for the mosquitoes that spread WNV, the presence of highly efficient vectors of WNV, or a local concentration of a population of highly susceptible horses (eg, farms with inadequate vaccination or management practices).

The other hot spot identified in eastern Texas could be explained by the shifting location of horses reported with WNV disease in Texas, a higher population at risk (a mean population of 2,111 horses/county) and therefore reporting bias, and suitable environmental factors that favor the vectors of WNV. In 2003, 61 horses with WNV disease were reported in Texas, of which only 1 affected horse was reported from the northwestern hot spot identified in this study. However in 2006, 111 horses with WNV disease were reported, of which 10 (9%) affected horses were reported from the eastern hot spot identified in this study. Based on the population at risk in this eastern hot spot (8,443 horses) and the total estimated number of horses in Texas (372,341 horses), we would only have expected 2.5 affected horses to be reported from this hot spot, if disease was evenly distributed in Texas. Thus, the eastern hot spot identified might represent a permanent area of high risk for WNV disease. It is of interest that this hot spot, as determined from maps of smoothed county attack rates, apparently became more pronounced during the period of 2002 to 2004. Study of environmental or management factors in this hot spot might reveal conditions that can lead to hyperendemic transmission of WNV in horse populations.

Results of our study indicate that locations of WNV disease in horses changed during the period of 2002 to 2004 and that this change was not substantially influenced by when reports of affected horses occurred within each of the 3 study years. In 2002, the first year WNV disease was reported in Texas, the mean center of affected horses was located in north-central Texas. By 2003, the mean center of reported horses had moved approximately 104 km southeast, and by 2004, the mean center of reported horses was located in eastern Texas, approximately 340 and 232 km southeast of the 2002 and 2003 mean centers, respectively.

A similar southeasterly pattern of progression was observed in the risk maps produced. This change in distribution might be a reflection of the way in which WNV was introduced into Texas and local environmental conditions. A large number of horses with WNV disease were reported from northwestern Texas during 2002. However, during 2003 and 2004 few affected horses were reported, with fewer counties reporting WNV disease in horses. Affected horses were consistently reported from eastern Texas during the period of 2002 to 2004. Thus, the large epidemic in northwestern Texas may have been an anomaly as a result of WNV introduction from north-central areas of the United States in 2002 (possibly via infected birds migrating south), whereas the environmental conditions (sources of water, vegetation types, and higher precipitation and warmer temperatures) and a larger population at risk in eastern Texas might be suitable for WNV endemicity.

The influence of vaccination is another issue that needs to be investigated. A killed vaccine to protect horses against WNV disease was available in 2002, prior to reports of horses with WNV disease in Texas. Shifts in the distribution of reported horses might have been caused by the adoption of vaccination as a management practice in different regions of Texas. For example, if horse owners in northwestern Texas were more likely to have vaccinated their horses in 2003 and 2004 as a result of experiencing the large epidemic in 2002, this might have caused the epicenters to appear to move southeast in Texas during 2003 and 2004.

Although the overall distributions of the 2002, 2003, and 2004 epidemics moved in a southeasterly direction across Texas, the distribution of each of the epidemics tended to be closer to the north, earlier in the season, affected horses were often reported from southern districts, and later in the season, affected horses were more often reported from northern districts. The most likely explanation of this pattern is a climate-driven process, specifically temperature. The activity of mosquitoes is known to be temperature dependent, and temperature-dependent models have been developed in an attempt to explain the spread of WNV disease within districts. It is biologically plausible that districts at lower latitudes would have a greater WNV-disease risk earlier in the season than districts at more northern latitudes. Although this pattern was observed in the data, the general direction of the occurrence of horses with WNV disease reveals variability. At a local scale, many other factors may be important in determining when disease risk is greatest.

The present study included reports of horses with WNV disease from 2002, 2003, and 2004. Inclusion of a longer period (eg, 2002 to 2006) might have provided some additional information on the pattern of reported horses with WNV disease in Texas. However, reporting bias is likely to have become an increasingly important issue in studies with data from latter years, as WNV became endemic in Texas. In 2005 and 2006, only 61 and 111 horses with WNV disease were reported, respectively. As horse owners learn to live with the disease, information from passive surveillance systems will have less use for research into the spatial distribution of, and risk factors for, WNV-disease occurrence.

In conclusion, hot spots of WNV disease developed within the horse population in Texas during the period of 2002 to 2004. The hot spot identified in the northwest area of the state might have been an anomaly related to introduction of WNV to Texas in 2002, but the hot spot identified in southeastern Texas appears more persistent. Determining factors associated with the hot spots could help in the development of more
effective disease control strategies, and locating specific hot spots could allow targeted disease prevention strategies to be implemented.

References