Human Plague in the Southwestern United States, 1957–2004: Spatial Models of Elevated Risk of Human Exposure to *Yersinia pestis*

REBECCA J. EISEN,[1,2] RUSSELL E. ENSCORE,[1] BRAD J. BIGGERSTAFF,[1] PAMELA J. REYNOLDS,[3] PAUL ETTESTAD,[3] TED BROWN,[4] JOHN PAPE,[5] DALE TANDA,[5] CRAIG E. LEVY,[6] DAVID M. ENGELTHALER,[6] JAMES CHEEK,[7] RUDY BUENO, JR.,[8] JOSEPH TARGHETTA,[8] JOHN A. MONTENIERI,[1] AND KENNETH L. GAGE[1]

J. Med. Entomol. 44(3): 530–537 (2007)

ABSTRACT Plague is a rare but highly virulent flea-borne zoonotic disease caused by the Gram-negative bacterium *Yersinia pestis* Yersin. Identifying areas at high risk of human exposure to the etiological agent of plague could provide a useful tool for targeting limited public health resources and reduce the likelihood of misdiagnosis by raising awareness of the disease. We created logistic regression models to identify landscape features associated with areas where humans have acquired plague from 1957 to 2004 in the four-corners region of the United States (Arizona, Colorado, New Mexico, and Utah), and we extrapolated those models within a geographical information system to predict where plague cases are likely to occur within the southwestern United States disease focus. The probability of an area being classified as high-risk plague habitat increased with elevation up to ≈2,300 m and declined as elevation increased thereafter, and declined with distance from key habitat types (e.g., southern Rocky Mountain piñon—juniper [*Pinus edulis* Engelm. and *Juniperus* spp.], Colorado plateau piñon–juniper woodland, Rocky Mountain ponderosa pine (*Pinus ponderosa* P.& C. Lawson var. scopulorum), and southern Rocky Mountain juniper woodland and savanna). The overall accuracy of the model was >82%. Our most conservative model predicted that 14.4% of the four-corners region represented a high risk of peridomestic exposure to *Y. pestis*.

KEY WORDS plague, *Yersina pestis*, fleas, geographical information system, landscape features

Plague, a primarily flea-borne zoonotic disease caused by the Gram-negative bacterium *Yersinia pestis* Yersin, was first reported from the United States in 1900. The initial cases occurred in San Francisco, and the pathogen is thought to have been introduced to this area by infected rats and fleas originating from ships arriving from plague-endemic regions of Asia (Eskey and Haas 1940, Pollitzer 1954, Gage and Kosoy 2005). It is likely that the pathogen persisted in urban rat populations in the San Francisco Bay area for some time before establishing sylvatic cycles involving California ground squirrels, *Spermophilus beechyii* (Richardson), and other rodent species (Eskey and Haas 1940, Meyer 1942, Poland and Barnes 1979, Barnes 1982). Despite control efforts, during the last century *Y. pestis* spread in rodent populations beyond the original focus in California and had virtually reached its current eastern limit (North Dakota, South Dakota, Kansas, Nebraska, Oklahoma, and Texas) before 1960 (Eskey and Haas 1940, Barnes 1982). Although widespread across the western United States, most human cases occur in the four-corners region (83% of 416 total cases from 1957 to 2004), which includes the states of Arizona, Colorado, New Mexico, and Utah.

Plague is characterized by quiescent and epizootic periods. The majority of human cases are thought to occur during epizootics, when susceptible hosts (typically ground squirrels, prairie dogs, chipmunks or woodrats) perish in large numbers, and infected fleas are forced to parasitize hosts upon which they would not ordinarily feed (Eskey and Haas 1940, Kartman et al. 1958, Barnes 1982, Gage et al. 1995, Levy and Gage 1999). Although it is a rare disease of humans in the United States, with only 112 cases reported during the 15-yr period between 1988 and 2002, fatality rates remain high (MMWR 2002). Mortality rates for untreated plague infections range from 50 to 60% for bubonic to nearly 100% for pneumonic or septicemic

[1] Division of Vector-Borne Infectious Diseases, National Center for Infectious Diseases, Centers for Disease Control and Prevention, P.O. Box 2087, Fort Collins, CO 80522.
[2] Corresponding author, e-mail: dyn2@cdc.gov.
[3] Zoonoses Program, New Mexico Department of Health, Santa Fe, NM 87502.
[4] Vector Control Program, New Mexico Environment Department, Santa Fe, NM 87502.
[5] Colorado Department of Health and the Environment, Denver, CO 80246.
[6] Vector-Borne and Zoonotic Disease Section, Arizona Department of Health Services, Phoenix, AZ 85007.
[7] Office of Environmental Health and Engineering, Navajo Area, Indian Health Service, Window Rock, AZ 86515.
[8] City of Albuquerque Division of Environmental Health, Albuquerque, NM 87102.

infection (Poland and Barnes 1979, Levy and Gage 1999). Outcome of infection is improved by early diagnosis and appropriate antibiotic treatment (Levy and Gage 1999; MMWR 2002, 2006; Gage and Kosoy 2005). Thus, the ability to predict when and where infections are likely to occur would be a useful tool to increase the local awareness of plague risk factors in the medical community and hence reduce the potential for plague-induced human mortality. It also improves our ability to target limited prevention resources.

Previous studies have identified climatological factors associated with increased plague activity. Typically, in the southwestern United States, epizootic activity intensifies when cool summer temperatures follow wet winters (Parmenter et al. 1999, Enscore et al. 2002). Similarly, epizootic activity in black-tailed prairie dogs of the north central plains of Colorado increased after El Niño events (Stapp et al. 2004). In the current study, we aimed to identify landscape features associated with human plague cases within the southwestern plague focus and to create a predictive geographical model of high-risk habitats for human exposure to *Y. pestis* in this region.

Materials and Methods

Study Area. Our study focuses on the four-corners region of the United States (Fig. 1, inset). The region is ecologically highly diverse. The short-grass steppe communities that dominate eastern Colorado and northeastern New Mexico give way to forests characterized by ponderosa pine, *Pinus ponderosa* Lawson; Douglas fir, *Pseudotsuga menziesii* Franco; and aspen (*Populus* spp.) in the northern part of the Rocky Mountains and to Colorado plateau shrublands further south. The latter habitat is characterized by piñon-juniper (*Pinus edulis* Engelm. and *Juniperus* spp.) and sagebrush at lower elevations and transitions into ponderosa pine; lodgepole pine, *Pinus contorta* Dougl.; and aspen as elevation increases (Dinerstein et al. 1999). Mammalian species richness in this region is among the highest found in North America (Findley 1987). The Colorado plateau shrublands are bounded by the dry Wasatch Uinta montane forest in north central Utah and the Arizona mountain forests of central Arizona. The Wasatch Uinta forests are dominated by Ponderosa pine, whereas the Arizona mountain forests transition from piñon–juniper Gambel oak, *Quercus gambelii* Nuttall, woodlands at low elevation to primarily pine forest at high elevation. South and west of these mountain ranges, desert communities dominate the landscape (Dinerstein et al. 1999).

The human population is focused in major cities located along the Front Range of central Colorado (e.g., Denver, Colorado Springs) and in north central New Mexico (Albuquerque and Santa Fe), southern Arizona (Phoenix, Tucson), or northern Utah (Salt Lake City). Nevertheless, the majority of human plague cases occur in rural and semirural areas in this region.

Construction of the Model Build and Validation Sets. *Case Points.* In total, 346 human cases of plague were reported to the U.S. Centers for Disease Control and Prevention from Arizona, Colorado, New Mexico, or Utah during the period 1957 through 2004. Based on information from epidemiological investigations conducted at the times these cases occurred, specific locations of exposure to *Y. pestis* could be determined for 316 of the cases. Peridomestic exposure, occurring within 2 km of a residence, was ascertained for 266 cases (84%). We created two build sets for the generation of two separate models. The first incorporates all 316 cases and focuses on risk habitats associated with peridomestic or nonperidomestic (e.g., recreation, occupational, or other) exposure. The second includes only peridomestic cases, for which the accuracy in identifying the site of exposure is likely more precise. Based on case investigations, latitude and longitude of probable sites of exposure were recorded, and these points were imported into a geographical information system (GIS). In total, 80% of case points ($n = 213$ for the peridomestic model; $n = 253$ for the peridomestic and recreational model) were used in the model build set. The remaining 20% of cases were reserved for model evaluation ($n = 53$ for the peridomestic model; $n = 63$ for the peridomestic and recreational model).

Control Points. Because plague is a severe and reportable disease, we assume that all cases of plague occurring in the region were included in the data set. Furthermore, with the exception of the 30 cases for which site of exposure could not be determined, all individuals in the population are assumed to be uninfected. Because we are modeling human cases of plague in relation to environmental characteristics, the random control points represent a random sampling of the human population rather than of the landscape. Because exact locations of control home sites were not available, we generated a control set that approximates the distribution of humans across the four corners landscape. To generate a random sample of 316 control points, we created a 30- by 30-m grid layer based 1990 Census block group data (population per square kilometer). Population density associated with the case points ranged from 0.23 to 151.43 people per km^2 in Arizona ($n = 56$ cases), 0.08–2,204.9 for Colorado ($n = 48$), 0.04–505.48 for New Mexico ($n = 208$), and 0.08–1.08 for Utah ($n = 4$). For each state, we drew a random, population density-weighted, control sample equal to the number of cases per state from a similar population range of cases per state. To avoid selecting control points from within the same grid cell, the minimum distance between control points was 1 km. Control points ($n = 213$) for the peridomestic model were randomly selected as a subset of the control points selected for the peridomestic and nonperidomestic model.

Predictive Landscape Features. Evaluated landscape features included elevation and habitat type. Elevation was derived from a one km resolution digital elevation model (U.S. Geological Survey [USGS]).

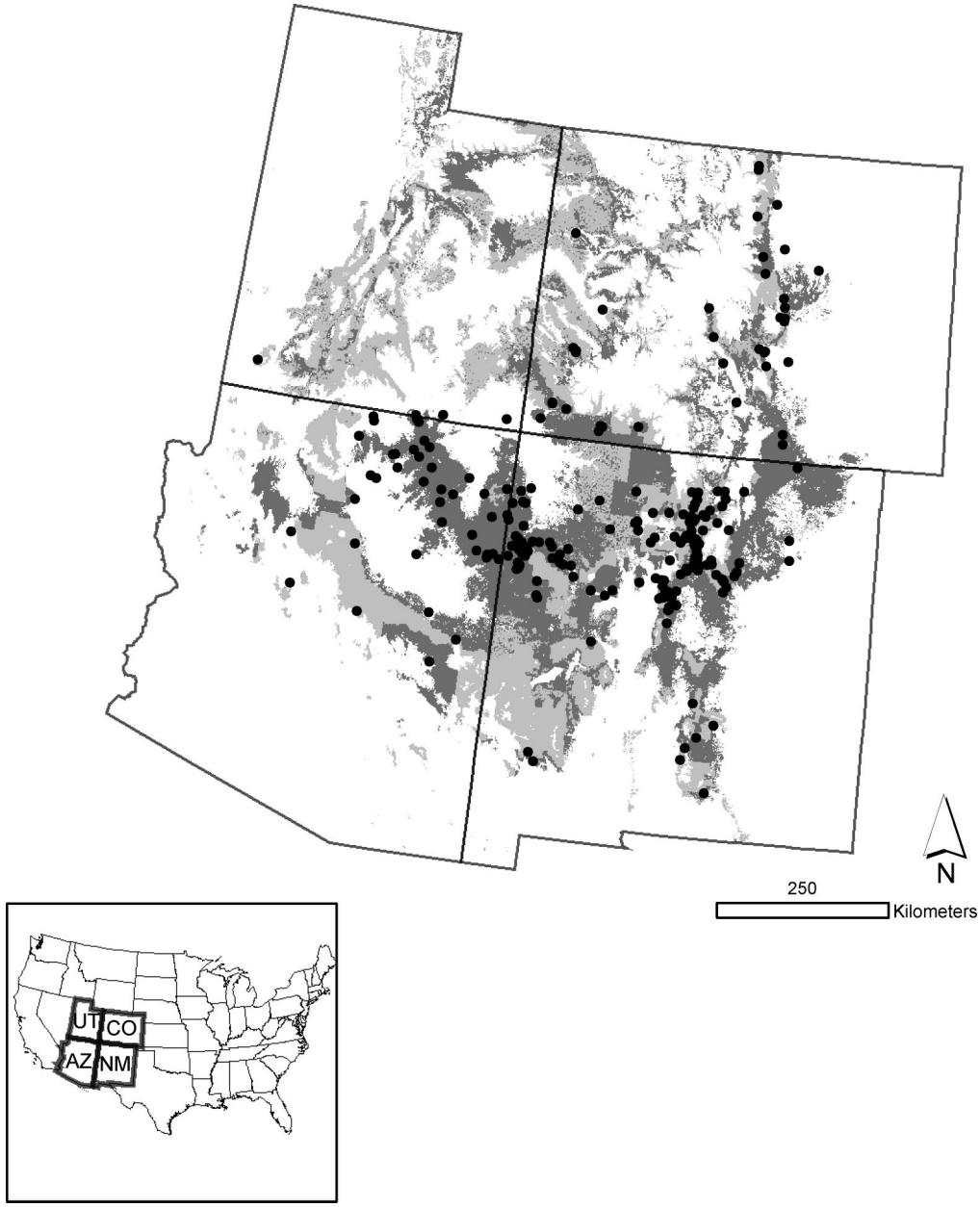

Fig. 1. Areas predicted by the peridomestic model to pose high risk to humans throughout the four-corners region (Arizona, Colorado, New Mexico, and Utah) are depicted in light gray. Those high-risk areas on privately or tribally owned land are shown in dark gray. Black circles represent locations of peridomestically acquired human plague cases from 1957 to 2004 in the four-corners region. States comprising the four corners region are shown within the United States in the inset.

Habitat type was determined based on classification of the USGS southwest regional Gap (30 m) land cover analysis (USGS 2004). Visualizing human case points in relation to habitat types revealed that four habitat types seemed to be associated with human plague cases. These included southern Rocky Mountain piñon–juniper, CO plateau piñon–juniper woodland, Rocky Mountain Ponderosa pine, and southern Rocky Mountain juniper woodland and savanna. These habitat types were described in detail previously (USGS 2004).

We generated 30- by 30-m grid layers representing the minimum Euclidean distance to any of these four habitats (hereinafter referred to as distance to key habitat), or minimum Euclidean distance to either southern Rocky Mountain piñon–juniper or Colorado plateau piñon–juniper woodland habitats (referred to as distance to piñon–juniper). All layers, including the

Table 1. Candidate models for associations of landscape features with human plague cases acquired either in the peridomestic environment or during peridomestic or recreational activities in the four corners region of the southwestern United States

Model no.	k[a]	Negative log-likelihood	AIC value[b]			Goodness of fit; *P* value	AUC[c]	Independent model variables[d]
			AIC	ΔAIC	Wt			
Peridomestic and nonperidomestic exposure combined								
1	4	252.15	512.30	0	0.9	0.35	0.81	DistKH, EL, EL^2
2	2	256.39	516.78	4.48	0.1	0.67	0.81	DistKH
3	4	261.98	531.96	19.66	0	0.11	0.80	DistPJ, EL, EL^2
4	3	263.79	533.58	21.28	0	0.11	0.80	EL, EL^2
5	2	285.10	574.20	61.90	0	0.06	0.80	DistPJ
Peridomestic exposure								
1	4	185.90	193.90	0	0.51	0.94	0.86	DistKH, EL, EL^2
2	2	190.07	194.07	0.17	0.47	0.99	0.84	DistKH
3	4	195.29	203.29	9.39	0.01	0.71	0.84	DistPJ, EL, EL^2
4	3	196.72	202.72	8.82	0.01	0.86	0.84	EL, EL^2
5	2	220.98	224.98	31.08	0	0.56	0.84	DistPJ

[a] k is number of estimated parameters included in the model.
[b] AIC, Akaike Information Criterion value; ΔAIC, AIC of model − AIC of best model.
[c] AUC, area under ROC curve.
[d] DistKH, distance to key habitat; DistPJ, distance to piñon–juniper habitat; EL, elevation; EL^2, elevation squared.

case and control points, were projected to NAD 1983 Albers.

Construction and Selection of Landscape Models. Logistic regression models were constructed to evaluate the association between the probability that an area is suitable plague habitat and its landscape features. Five candidate models were constructed for each of the peridomestic and nonperidomestic or peridomestic build sets (n = 506 or 426 case and control points included per model, respectively) (Table 1). The models may be described by the following equation:

$$\text{Logit}(P) = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \ldots + \beta_k x_k \quad [1]$$

where P is the probability of a grid cell being classified as suitable plague habitat, and β_0 is the intercept. The values $\beta_{1,\ldots,}\beta_k$ represent the coefficients assigned to each independent variable, $x_1, \ldots, x_k$, included in the regression. The probability that a particular cell is classified as suitable habitat can be derived from equation 1 by using the following expression:

$$P = \exp(\beta_0 + \beta_1 x_1 + \ldots + \beta_k x_k)/[1 + \exp(\beta_0 + \beta_1 x_1 + \ldots + \beta_k x_k)] \quad [2]$$

To select the most parsimonious model with high predictive power, Akaike's Information Criterion (AIC) (Akaike 1974) was used to rank each of the models. The one with the lowest AIC value was selected. However, models within two AIC units of the minimum AIC ($\Delta AIC < 2$) are considered competing with substantial support (Burnham and Anderson 1988). To determine the amount of evidence in favor of a particular model, we calculated Akaike weights (ω_i) for each model i = 1, 2, . . . , 5. We also used a goodness-of-fit test to determine whether the covariates included in the model adequately explained the distribution in the data. This analysis compares a pure-error negative log-likelihood with the fitted model log-likelihood. If the chi-square test is not significant, it supports the conclusion that sufficient explanatory variables were included in the model.

Receiver operating characteristic curves (ROCs) were used to assess the overall discrimination ability of each model, based on the area under the ROC curve (AUC), and to determine the optimal probability cut-off for characterization of a grid cell as either high- or low-risk plague habitat. An ROC curve plots all true positive fractions (sensitivity values) obtained from the model build set on the vertical axis against their corresponding equivalent false positive fraction values (1 − specificity) for all available thresholds on the horizontal axis. The AUC provides a threshold-independent measure of the overall accuracy of the model. This value ranges from 0.5 to 1, where a value of one indicates that all points in the build set were correctly classified by the model (Fielding and Bell 1997). The logit equation can be transformed into a probability following equation 2. However, to evaluate the performance of the model by using ROC, this probability is converted to a binary value (e.g., a grid cell is either classified as high plague risk or low plague risk habitat). The optimal probability cut-off value was chosen by maximizing sensitivity and specificity simultaneously (Fielding and Bell 1997, Guisan and Zimmerman 2000). All cells with a probability value equal to or greater than the optimal value were classified as high-risk plague habitat. All others were considered low risk in our evaluation matrix.

Results

Summary of Human Case Data. Site of exposure to the etiological agent of plague could be determined for 316 (91.3%) of the 346 human cases occurring from 1957 to 2004 in the four-corners region. Among these cases, 208 (65.83%) were derived from New Mexico, 56 (17.72%) from Arizona, 48 (15.19%) from Colorado, and four (1.27%) from Utah. For all 316 cases, 50%

were contracted at elevations between 1,856 and 2,178 m; of the remainder, 25% occurred below 1,856 m and 25% above 2,178 m. Forty-one (13%) of the 316 cases were fatal.

Mode of transmission could be determined for 80% ($n = 252$) of the 316 cases. Flea-borne transmission was most common (83.7%; $n = 211$), followed by direct contact (15.1%; $n = 38$) and air-borne exposure (1.2%; $n = 3$). Domestic cats served as the source of all of the air-borne exposures; infected cats can expel respiratory droplets infected with *Y. pestis* (Gage et al. 2000). The majority of cases associated with direct contact involved either rabbits (44.74%; $n = 17$) or domestic cats (26.32%; $n = 10$). Other sources of exposure related to direct contact included prairie dogs (15.79%; $n = 6$), bobcats (5.26%; $n = 2$), gray foxes (5.26%; $n = 2$), and a coyote (2.6%; $n = 1$). For flea-borne transmission, the vertebrate association of the infected flea(s) was unknown for 72 cases (34.12%). Among the remaining 139 flea-associated cases, fleas were most commonly reported to originate from rock squirrels (72.7%; $n = 101$), followed by prairie dogs (13.7%; $n = 19$), ground squirrels (7.9%; $n = 11$), rabbits (2.2%; $n = 3$), domestic cats (1.4%; $n = 2$), woodrats (1.4%; $n = 2$), and a chipmunk (0.7%; $n = 1$).

Peridomestically and Nonperidomestically Acquired Plague. *Association of Landscape Features with Human Plague.* Among the candidate logistic regression models, the most parsimonious model (Table 1) for predicting areas where human plague cases were reported included elevation (8.971×10^{-3} m^{-1} $\pm$ 3.2909×10^{-3}; $\beta \pm$ SE), elevation^2 (-2.0×10^{-6} m^{-2} $\pm$ 8.0×10^{-7}), and distance to key habitat (southern Rocky Mountain piñon–juniper, Colorado piñon–juniper, Rocky Mountain Ponderosa pine, or southern Rocky Mountain juniper woodland desert and savanna) (-3.225×10^{-4} m^{-1} $\pm$ 8.42×10^{-5}; intercept -8.80 ± 3.46). The model indicated that suitability of habitat for plague cases increases up to an elevation of $\approx$ 2,240 m and declines as elevation increases thereafter.

Case points occurred closer to the key habitats than controls (Wilcoxon ranked sums test: $\chi^2 = 178.1$, df = 1, $P < 0.0001$). Accuracy of the best model, based on the area under the ROC curve was 0.82. This value indicates that 82% of the time, randomly selected high- and low-risk pairs will be correctly ordered by their probability/habitat scores.

Modeling High-Risk Plague Habitat. Probability of suitability within the four-corners region was calculated based on the model described above. Using the optimal cut-off probability value, $P = 0.6409$, derived from the ROC curve, the model predicted that 35.14% of the area represents high-plague risk habitat. However, most human plague cases in this region are acquired peridomestically, rather than recreationally. Therefore, we calculated the percentage of the state that was classified by the federal land survey layer (U.S. National Atlas, U.S. Department of the Interior) as nonfederal or tribal ownership and had a model prediction probability of plague $\geq$0.6409. Under these restrictions, 18.8% of the four-corners region is considered high-risk habitat.

Table 2. Evaluation matrix, based on build set data, for the most parsimonious human plague case model for peridomestic and nonperidomestic exposure or peridomestic exposure

Model classification	Actual classification		% correct[b]
	Plague case	Control	
Peridomestic and nonperidomestic[a]			
High-plague risk	204	71	74.18
Low-plague risk	49	182	78.79
% correct[c]	80.63	71.94	
Peridomestic[d]			
High-plague risk	178	48	78.76
Low-plague risk	35	165	82.50
% correct[c]	83.57	77.46	

[a] Probability cut-off value used to classify high plague risk based on the ROC optimal cut-off probability ($P \geq 0.6409$).
[b] User accuracy (commission error).
[c] Producer accuracy (omission error).
[d] Probability cut-off value used to classify high plague risk based on the ROC optimal cut-off probability ($P \geq 0.6819$).

Model Evaluation. The model prediction and actual point classification (case or control) are presented in Table 2. Overall, user accuracies for correctly classifying high- and low-risk plague habitats, based on a probability cut-off greater than or equal to 0.6409, were 74.18 and 78.79%. This indicates that 74.18% of cells in the evaluation set classified by the model as high-risk habitat contained human case points. Errors of commission (classifying a control point as falling within high-risk habitat) occurred 25.82% of the time. In contrast, producer accuracies for correctly classifying high- and low-risk habitats were 80.63 and 71.94%, respectively. That is, 80.63% of actual case points were located in cells classified by the model as high plague-risk habitat, and 71.94% of control points were classified as low risk. Errors of omission, where a cell containing a case point was misclassified as low-risk plague habitat, were infrequent (19.37%).

Evaluation of the model by using an independent evaluation set ($n = 63$ cases and 63 controls) showed similar accuracies (Table 3). Overall, 82.5% of case points were located in cells classified as high-risk plague habitat and 73.2% of habitat classified by the model as high-risk contained case points. In total, 30.16% of control points were located in cells classified as low-plague risk habitat. This error is not unusual in situations where suitable habitat is not saturated with the species of interest that is being modeled (Fielding and Bell 1997), in this case a rare disease in which habitat may be suitable for plague to persist in a zoonotic cycle, but human exposure to the pathogen is low.

Peridomestically Acquired Plague. *Association of Landscape Features with Human Plague.*

Like the model built with peridomestic and nonperidomestic cases, the most parsimonious model (Table 1) for predicting areas where human plague cases were reported included elevation ($1.2 \times 10^{-2} \pm 4.7 \times 10^{-3}$ m^{-1}; $\beta \pm$ SE), elevation^2 (-2.7×10^{-6} m^{-2} $\pm$ 1.1×10^{-6}), distance to key habitat (southern Rocky Mountain piñon–juniper, Colorado piñon–juniper, Rocky Mountain Ponderosa pine, or southern Rocky

Table 3. Evaluation matrix, based on evaluation set data, for the most parsimonious human plague case model for peridomestic and nonperidomestic exposure or peridomestic exposure

Model classification	Actual classification		% correct[b]
	Plague case	Control	
Peridomestic and nonperidomestic[a]			
High-plague risk	52	19	73.24
Low-plague risk	11	44	80.00
% correct[c]	82.54	69.84	
Peridomestic[d]			
High-plague risk	42	16	72.41
Low-plague risk	11	37	77.08
% correct[c]	79.25	69.81	

[a] Probability cut-off value used to classify high plague risk based on the ROC optimal cut-off probability ($P \geq 0.6409$).
[b] User accuracy (commission error).
[c] Producer accuracy (omission error).
[d] Probability cut-off value used to classify high plague risk based on the ROC optimal cut-off probability ($P \geq 0.6819$).

Mountain juniper woodland desert and savanna) ($-6.4 \times 10^{-4}\ m^{-1} \pm 1.9 \times 10^{-4}$; intercept -12.36 ± 5.01). The model indicated that suitability of habitat for plague cases increases up to an elevation of ≈ 2,222 m and declines as elevation increases thereafter. Case points occurred closer to the key habitats than control points ($\chi^2 = 153.8$, df = 1, $P < 0.0001$). Accuracy of the best model based on the area under the ROC curve (0.86) was similar to the more inclusive model.

Modeling High-Risk Plague Habitat. Probability of suitability within the four-corners region was calculated based on the model described above. Using the optimal cut-off probability value, 0.6819, based on the ROC curve, the model predicted that 27.9% of the area is considered high plague-risk habitat (Fig. 1). Again, assessing risk for areas classified by the federal land survey layer as nonfederal or tribal ownership and had a model prediction probability of plague ≥0.6819, the area considered at high risk was considerably reduced. Under these conditions, 14.4% of the four corners region is considered high-risk habitat (New Mexico, 22.1%; Arizona, 13.0%; Colorado, 12.7%; and Utah, 7.5%) (Fig. 1).

Model Evaluation. The model prediction and actual point classification (case or control) are presented in Table 2. Overall, user accuracies for correctly classifying high- and low-risk plague habitats, based on a probability cut-off of $P \geq 0.6819$, were 78.8 and 82.5%, respectively. This indicates that 78.8% of cells included in the evaluation set and classified by the model as high-risk habitat contained a human case point. Errors of commission (classifying a control point as falling within high-risk habitat) occurred 21.2% of the time. Producer accuracies for correctly classifying high- and low-risk habitats were 83.57 and 77.46%, respectively. That is, 83.57% of actual case points were located in cells classified by the model as high plague-risk habitat, and 77.46% of control points were classified as low risk. Errors of omission, where case points were misclassified as low risk habitat, were infrequent (16.43%).

Evaluation of the model using an independent evaluation set ($n = 53$ cases and 53 controls) showed similar accuracy (Table 3). Overall, 79.25% of case points were located within cells classified as high-risk plague habitat, and 72.4% of habitat was classified by the model as high-risk contained case points. In total, 30.1% of control points were located in cells classified as high-risk plague habitat.

Discussion

Description and Distribution of High-Risk Plague Habitat in the Four-Corners Region of the Southwestern United States. Our model indicates that high-risk habitats can be classified accurately based on a nonlinear relationship with elevation and distance to key habitats. The most conservative model, which was based solely on peridomestic cases and extrapolated to nonfederal land, predicted that 14.4% of the four-corners region posed an elevated risk of peridomestic exposure. Among the four states, New Mexico has the highest coverage of peridomestic risk habitat (22.1%) and also reported the highest number of human cases (195 confirmed cases during 1957–2004). In contrast, Utah was classified as the state with the lowest coverage of peridomestic risk habitat (7.5%) and reported the fewest human cases (four cases). Although the percentage of the region considered at high risk of peridomestic exposure to infection is relatively low, some risk areas do coincide with areas where human population density is already high (e.g., Santa Fe County, NM) or increasing rapidly (Colorado Front Ranges). Increased human usage of these areas is likely to result in more frequent human contact with the infectious agent, and, consequently, rising plague incidence (Meyer 1955, Craven et al. 1993).

Rodent and Flea Communities Associated with Key Risk Habitats. We have identified key habitat types associated with the occurrence of human plague cases. We think these habitat types represent areas rich in rodent hosts that are highly susceptible to *Y. pestis* (epizootic hosts) and that live in close proximity to human habitations. These habitats also probably harbor large numbers of fleas that are competent vectors of the plague bacterium and are thought to bite humans when their natural hosts die.

In North America, it is generally accepted that plague persists in enzootic and epizootic cycles (Gage and Kosoy 2005). Enzootic (maintenance) cycles involve resistant host populations that have many members that are likely to survive plague, but they also have some members that suffer more severe illness and experience bacterial loads sufficient to infect at least a proportion of feeding fleas. Proposed enzootic hosts include deer mice, *Peromyscus maniculatus* Wagne; California meadow voles, *Microtus californicus* Peale; and northern grasshopper mice, *Onychomys leucogaster* Wied-Neuwied (Poland and Barnes 1979; Thomas 1988a, 1988b; Gage et al. 1995). It is presumed that epizootics occur when *Y. pestis* is introduced from enzootic cycles to highly susceptible epizootic host populations. Epizootic (amplifying) hosts include

ground squirrels (*Spermophilus* spp.), prairie dogs (*Cynomys* spp.), woodrats (*Neotoma* spp.), and chipmunks (*Tamias* spp.) (Holdenried and Morlan 1956, Holdenried and Quan 1956, Barnes 1982, Gage et al. 1995). Flea species infesting these enzootic and epizootic hosts are numerous and have been listed previously (Haas et al. 1973, Barnes 1982, Thomas et al. 1988b, Gage et al. 1995). Epizootic hosts are thought to play a key role in human plague epidemiology because the high bacterial loads required to infect feeding fleas (Wheeler and Douglas 1945, Burroughs 1947, Lorange et al. 2005) are typically observed in terminally septicemic hosts collected during epizootics. The subsequent death of these hosts forces potentially infected fleas to seek new hosts, including humans.

It has been proposed that epizootic activity is most likely to occur in areas where multiple populations of highly susceptible hosts are living at high densities in diverse and patchy habitats (Gage et al. 1995). These conditions can be found in mountainous and plateau regions of Arizona, Colorado, New Mexico, and Utah, which include the four key-risk habitat types included in our model. Rock squirrels, *Spermophilus variegates* Benson; woodrats (*Neotoma* spp.); chipmunks (*Tamias* spp.), and Gunnison's or white-tailed prairie dogs, *Cynomys gunnisoni* Baird and *Cynomys leucurus* Merriam, respectively, are known to occur within these habitat types, and each was implicated in at least one of the human cases used to construct our models (Eskey and Haas 1940, Holdenried and Morlan 1955, Holdenried and Morlan 1956, Haas et al. 1973, Findley et al. 1975, Burt and Grossenheider 1980, Findley 1987, Fitzgerald et al. 1994, Hoogland 1995).

Notably, rock squirrels directly contributed to the highest number of human infections in both this and previous studies and their geographical distribution mirrors our risk map (Craven et al. 1993). Rock squirrels commonly live in close proximity to humans and readily inhabit artificial refuges such as rock walls, uninhabited outbuildings, and refuse dumps (Fitzgerald et al. 1994). Intriguingly, their population size seems to increase with human development in piñon–juniper habitats (Barnes 1982). These rodents are highly susceptible to *Y. pestis* infection; suffer high mortality after an infectious septicemic period; are heavily infested with rock squirrel fleas, *Oropsylla montana* Baker, that have recently been demonstrated to be highly efficient vectors of *Y. pestis;* and in the absence of rodents, aggressively seek alternative hosts, including humans (McCoy 1910, Wheeler and Douglas 1945, Burroughs 1947, Barnes 1982, Eisen et al. 2006).

Factors Affecting Accuracy of the Model. Evaluation of the two models, using both build set data and independent evaluation sets, demonstrated high predictive power (Tables 2 and 3). Accuracy was reduced by one of two classification errors. Either a control point was located within an area classified by the model as high-risk plague habitat (false positives), or a plague case was located within an area classified as low risk (false negative). Because plague is a rare disease of humans that is dependent on contact rates between humans and an infectious agent whose presence in any given location is transient, the percentage of false positives seen in our models was expected. Local variation in where woodrats and rock squirrels nest in relation to where humans live, differences in human activity outdoors, and extent of human contact with free-roaming pets that could become infected or carry infectious fleas all affect the likelihood of becoming infected regardless of high- or low-risk classification (Mann et al. 1979, Barnes 1982, Levy and Gage 1999). These factors were not included in the construction of our models and probably explain many of the misclassification errors. Misclassification errors also could be related to selection of control points (Pearce and Boyce 2006). Creation of habitat suitability models is often difficult because absence cannot be equated with unsuitable. In some instances, controls could have been selected from areas considered suitable plague habitat, thus obscuring differences between high and low risk.

False negative classification errors could be reduced if we had set the probability cut-off at a value lower than the optimal value determined by the ROC curve. This would result in an increase in the total coverage area considered to be high-risk habitat and would increase the number of cases in high risk plague habitat (true positives) at the expense of increasing the number of controls located within high-risk plague habitat (false positives).

Implications for Public Health. Our most inclusive and best predictive model of high-risk areas for plague included both peridomestic and recreational risk and private as well as public land. This liberal model is suited for identifying high-risk habitats and thus should aid in raising plague awareness among the medical community and in allocating limited resources. In contrast, the most conservative model (Fig. 1) was restricted to peridomestic risk and excluded federally owned land where few people live. This conservative model should be particularly useful for targeting expensive prevention and control efforts in heavily populated, high-risk areas. Both models are based on factors that are easily quantified by the general public, public health officials, and health care providers; thus, they provide a useful tool for raising awareness of areas that potentially pose a risk of human exposure to the etiologic agent of plague.

Acknowledgments

We thank L. G. Carter for testing diagnostic samples submitted to the Center for Disease Control and Prevention and L. Eisen for comments on the manuscript.

References Cited

Akaike, H. 1974. A new look at the statistical model identification. IEEE Trans. Autom. Control 19: 716–723.

Barnes, A. M. 1982. Surveillance and control of bubonic plague in the United States. Symp. Zool. Soc. Lond. 50: 237–270.

Burnham, K. P., and D. R. Anderson. 1988. Model selection and inference: a practical information-theoretic approach. Springer, New York.

Burroughs, A. L. 1947. Sylvatic plague studies: the vector efficiency of nine species of fleas compared with *Xenopsylla cheopis.* J. Hyg. 43: 371–396.

Burt, W. H., and R. P. Grossenheider. 1980. A field guide to the mammals of North America north of Mexico. Houghton Mifflin, Boston, MA.

Craven, R. B., G. O. Maupin, M. L. Beard, T. J. Quan, and A. M. Barnes. 1993. Reported cases of human plague infections in the United States, 1970–1991. J. Med. Entomol. 30: 758–61.

Dinerstein, E., D. M. Olson, C. J. Loucks, W. Eichbaum, and T. H. Ricketts. 1999. Terrestrial ecoregions of North America: a conservation assessment. Island Press, Washington, DC.

Eisen, R. J., S. W. Bearden, A. P. Wilder, J. A. Montenieri, M. F. Antolin, and K. L. Gage. 2006. Early-phase transmission of *Yersinia pestis* by unblocked fleas as a mechanism explaining rapidly spreading plague epizootics. Proc. Nat. Acad. Sci. U.S.A. 103: 15380–15385.

Enscore, R. E., B. J. Biggerstaff, T. L. Brown, R. F. Fulgham, P. J. Reynolds, D. M. Engelthaler, C. E. Levy, R. R. Parmenter, J. A. Montenieri, J. E. Cheek, et al. 2002. Modeling relationships between climate and the frequency of human plague cases in the southwestern United States, 1960–1997. Am. J. Trop. Med. Hyg. 66: 186–196.

Eskey, C. R., and V. H. Haas. 1940. Plague in the western part of the United States. Publ. Health Bull. 254: 1–83.

Fielding, A. H., and J. F. Bell. 1997. A review of methods for the assessment of prediction errors in conservation presence/absence models. Environ. Conserv. 24: 38–49.

Findley, J. S. 1987. The natural history of New Mexican mammals. The University of New Mexico Press, Albuquerque, NM.

Findley, J. S., A. H. Haris, D. E. Wilson, and C. Jones. 1975. Mammals of New Mexico. University of New Mexico Press, Albuquerque, NM.

Fitzgerald, J. P., C. A. Meaney, and D. M. Armstrong. 1994. Mammals of Colorado. University Press of Colorado, Niwot, CO.

Gage, K. L., and M. Y. Kosoy. 2005. Natural history of plague: perspectives from more than a century of research. Annu. Rev. Entomol. 50: 505–528.

Gage, K. L., R. S. Ostfeld, and J. G. Olson. 1995. Nonviral vector-borne zoonoses associated with mammals in the United States. J. Mammal. 76: 695–715.

Gage, K. L., D. T. Dennis, K. A. Orloski, P. Ettestad, T. L. Brown, P. J. Reynolds, W. J. Pape, C. L. Fritz, L. G. Carter, and J. D. Stein. 2000. Cases of cat-associated human plague in the Western US, 1977–1998. Clin. Infect. Dis. 30: 893–900.

Guisan, A., and N. E. Zimmerman. 2000. Predictive habitat distribution models in ecology. Ecol. Model. 135: 147–186.

Haas, G. E., R. P. Martin, M. Swickard, and B. E. Miller. 1973. Siphonaptera-mammal relationships in northcentral New Mexico. J. Med. Entomol. 10: 281–289.

Holdenried, R., and H. B. Morlan. 1955. Plague-infected fleas from northern New Mexico wild rodents. J. Infect. Dis. 96: 133–137.

Holdenried, R., and H. B. Morlan. 1956. A field study of wild mammals and fleas of Santa Fe County, New Mexico. Am. Midl. Nat. 55: 369–381.

Holdenried, R., and S. F. Quan. 1956. Susceptibility of New Mexico rodents to experimental plague. Publ. Health Rep. 71: 979–984.

Hoogland, J. L. 1995. The black-tailed prairie dog: social life of a burrowing mammal. The University of Chicago Press, Chicago, IL.

Kartman, L., F. M. Prince, S. F. Quan, and H. E. Stark. 1958. New knowledge on the ecology of sylvatic plague. Ann. N.Y. Acad. Sci. 70: 668–711.

Levy, C. E., and K. L. Gage. 1999. Plague in the United States, 1995–1997. Infect. Med. 16: 54–64.

Lorange, E. A., B. L. Race, F. Sebbane, and B. Joseph Hinnebusch. 2005. Poor vector competence of fleas and the evolution of hypervirulence in *Yersinia pestis.* J. Infect. Dis. 191: 1907–1912.

Mann, J. M., W. J. Martone, J. M. Boyce, A. F. Kaufmann, A. M. Barnes, and N. S. Weber. 1979. Endemic human plague in New Mexico: risk factors associated with infection. J. Infect. Dis. 140: 397–401.

McCoy, G. W. 1910. A note on squirrel fleas as plague carriers. Publ. Health Rep. 25: 465.

Meyer, K. F. 1942. The ecology of plague. Medicine 21: 143–174.

Meyer, K. F. 1955. The oern outlook on plague in California. Vector Views 2: 41–43.

[MMWR] Morbidity and Mortality Weekly Report. 2002. Imported plague: New York City, 2002. Morb. Mort. Wkly. Rep. 52: 725–728.

[MMWR] Morbidity and Mortality Weekly Report. 2006. Human plague-four states, 2006. Morb. Mort. Wkly. Rep. 55: 1–3.

Parmenter, R. R., E. P. Yadav, C. A. Parmenter, P. Ettestad, and K. L. Gage. 1999. Incidence of plague associated with increased winter-spring precipitation in New Mexico. Am. J. Trop. Med. Hyg. 61: 814–821.

Pearce, J. L., and M. S. Boyce. 2006. Modelling distribution and abundance with presence-only data. J. Appl. Ecol. 43: 405–412.

Poland, J. D., and A. M. Barnes. 1979. Plague, pp. 515–559. *In* J. H. Steele [ed.], CRC handbook series in zoonoses. Section A: bacterial, rickettsial and mycotic diseases, volume I. CRC, Boca Raton, FL.

Pollitzer, R. 1954. Plague. W.H.O. Monogr. Ser. No. 22. World Health Organization, Geneva, Switzerland.

Stapp, P., M. F. Antolin, and M. Ball. 2004. Patterns of extinction in prairie dog metapopulations: plague outbreaks follow El Niño events. Front. Ecol. 2: 235–240.

Thomas, R. E. 1988a. A review of flea collection records from *Onychomys leucogaster* with observations on the role of grasshopper mice in the epizoology of wild rodent plague. Great Basin Nat. 48: 83–95.

Thomas, R. E., A. M. Barnes, T. J. Quan, M. L. Beard, L. G. Carter, and C. E. Hopla. 1988b. Susceptibility to *Yersinia pestis* in the northern grasshopper mouse (*Onychomys leucogaster*). J. Wildl. Dis. 24: 327–333.

[USGS] U.S. Geological Survey. 2004. Southwest regional Gap Analysis project final report. Utah State University and United States Geological Survey Gap Analysis program, Logan, UT.

Wheeler, C. M., and J. R. Douglas. 1945. Sylvatic plague studies. V. The determination of vector efficiency. J. Infect. Dis. 77: 1–12.

Received 27 November 2006; accepted 24 January 2007.