

Relationship of Climate, Geography, and Geology to the Incidence of Rift Valley Fever in Kenya during the 2006–2007 Outbreak

Allen Hightower,* Carl Kinkade, Patrick M. Nguku, Amwayi Anyangu, David Mutonga, Jared Omolo, M. Kariuki Njenga, Daniel R. Feikin, David Schnabel, Maurice Ombok, and Robert F. Breiman

Division of Parasitic Diseases and Malaria, Center for Global Health, Center for Disease Control, Atlanta, Georgia; Epidemiology and Analysis Program Office, Office of Surveillance Epidemiology and Laboratory Systems, Center for Disease Control, Atlanta, Georgia; Division of Disease Surveillance and Response, Ministry of Public Health and Sanitation, Nairobi, Kenya; Field Epidemiology and Laboratory Training Program, Ministry of Public Health and Sanitation, Nairobi, Kenya; United States Army Medical Research Unit–Kenya, Nairobi, Kenya; Center for Global Health Research, Kenya Medical Research Institute, Kisumu, Kenya; Global Disease Detection Division, Centers for Disease Control and Prevention–Kenya, Nairobi, Kenya

Abstract. We estimated Rift Valley fever (RVF) incidence as a function of geological, geographical, and climatological factors during the 2006–2007 RVF epidemic in Kenya. Location information was obtained for 214 of 340 (63%) confirmed and probable RVF cases that occurred during an outbreak from November 1, 2006 to February 28, 2007. Locations with subtypes of solonetz, calcisols, solonchaks, and planosols soil types were highly associated with RVF occurrence during the outbreak period. Increased rainfall and higher greenness measures before the outbreak were associated with increased risk. RVF was more likely to occur on plains, in densely bushed areas, at lower elevations, and in the Somalia acacia ecological zone. Cases occurred in three spatial temporal clusters that differed by the date of associated rainfall, soil type, and land usage.

INTRODUCTION

Rift Valley fever (RVF) is a mosquito-borne viral zoonosis that causes periodic epidemics and epizootics in sub-Saharan Africa.¹ Although outbreaks are often associated with heavy rainfall and flooding² and have reoccurred in similar locations,^{3,4} heavy rainfall and flooding contribute to but are not the sole environmental criteria for RVF outbreaks.

RVF virus is transmitted by various species of mosquitoes, such as *Aedes* and *Culex*, and through the secretions of infected animals, which are also infected by the same range of mosquito vectors.^{5–9} Thus, the ecological niche of these mosquito species can help define the locations in which human cases might occur. The ecological niche of RVF vectors varies widely.¹ However, in eastern Africa, outbreaks have occurred in areas where drought-resistant vectors lay eggs that can survive for several years and require flooding events for hatching.¹ Additionally, the distance that the livestock is moved is often reasonably small, because much of the livestock is moved in this region by foot. Inclement weather conditions also limit the mobility of livestock populations during epidemic periods.¹⁰ Although considerable work has been conducted to predict outbreaks of RVF,^{11–14} the possible impact of geographical and geological factors in addition to climatological influences on the incidence of RVF disease in an area have not been characterized.

In this paper, we use geocoded case locations and their geographical, geological, and climatological attributes to estimate incidence of RVF disease during the outbreak period in Kenya in late 2006 to early 2007. We present both national models of person-time–based incidence covering the entire outbreak period and separate analyses for the three major temporal–spatial clusters. We use sources of climate, geographical, and geological data that are freely available on the

Internet and are available for many countries where RVF cases and outbreaks occur. The findings from these models could be used in outbreak prediction models to reduce the number of false outbreak predictions and serve as a basis for additional investigation for case occurrence in other locations.

METHODS

Case definition and case detection. A suspect case of RVF was defined as a person presenting between November 1, 2006 and February 28, 2007, with an acute febrile illness ($> 37.5^{\circ}\text{C}$ for > 48 hours) and not responding to antimicrobial drugs or antimalarial therapy in a district where human or livestock RVF was confirmed.⁴ A probable case was defined as a patient with fever and bleeding manifestations. The RVF cases were reported through the Kenya Ministry of Public Health and Sanitation's Integrated Disease Surveillance and Response (IDSR), a passive system that is used by most sub-Saharan African nations to monitor and control priority communicable and non-communicable disease. Patients meeting the probable or suspect case definition were defined as confirmed cases if immunoglobulin M (IgM) antibodies to RVF were detected by enzyme-linked immunosorbent assay (EIA) and/or RVF ribonucleic acid (RNA) was detected by reverse-transcriptase polymerase chain reaction (PCR). For the purpose of this report, all suspect patients (not confirmed by laboratory diagnostics) as well as probable cases from which specimens were negative by laboratory testing for RVF were excluded. Confirmed cases, as well as probable cases without available specimens who died before specimens could be obtained or who did not have access to healthcare during their acute illness (usually because of the widespread flooding), were included.⁴ Confirmed and probable cases without a geographic location at least at the division level (third administrative level) were excluded.

Data sources. Data layers for geologic, geographic, and demographic data, including soil types and land use patterns, were obtained from the publicly available geographic information system (GIS) site of the International Livestock Research

*Address correspondence to Allen Hightower (retired), Division of Parasitic Diseases and Malaria, Center for Global Health, Centers for Disease Control and Prevention, Mailstop A06, 1600 Clifton Road, Atlanta, GA 30333. E-mail: awh1953@gmail.com

Institute (ILRI).¹⁵ Soil-type data on the ILRI website were obtained from the Kenya Soil Survey, Report E1, 1982. Soils were classified by physical and chemical properties using the Food and Agriculture Organization (FAO) scheme.¹⁶ The land use data were obtained from Land-Use Satellite (LANDSAT) images obtained in 1987 by Japan International Co-operation Agency (JICA) for the Kenya National Water Master Plan. The ecological zone data were also obtained from the ILRI website. The Somali acacia ecozone is characterized by short grasses, shrubs, and acacia trees that can survive for extended periods without water.¹⁷ The 250-m digital elevation model data file was obtained from the World Resources Institute website.¹⁸

Normalized difference vegetation indices (NDVIs) and rainfall maps for 10-day periods from November 1, 2006 to February 28, 2007 were obtained from the Africa Data Dissemination Service.¹⁹ The NDVI is a commonly used measure of remotely sensed vegetation cover or greenness. It has a range from -1 , where there is little or no vegetation (and therefore, greenness), to $+1$, which corresponds to intense greenness. In areas where there are distinct rainy and dry seasons, the NDVI often increases after the rainy season begins, and therefore, the 10-day NDVI can be more closely related to a 10-day rainfall measure in the recent past. As a result, NDVI is considered to be a lagging indicator of rainfall in the past month. For longer periods of rainfall, vegetation levels often level off, decreasing the relationship between rainfall and NDVI.²⁰

Two types of rainfall measures were used. One was an estimated annual rainfall in centimeters.¹⁵ The second was a series of rainfall estimates for each 10-day period (called dekads) from November 1, 2006 to February 28, 2007.¹⁹ Population and population density data for sublocations were based on the 1999 Kenya Census.²¹ The sublocation is the smallest administrative unit in Kenya. There were 6,625 sublocations in 1999. The population was considered to be at risk from November 1, 2006 to February 28, 2007.

GIS methods. To create the analysis database of grid cells with the associated geographic information, the country was divided into a grid containing 46,200 cells (or squares) with a length of 3.5 km on each side. Each of these cells became the basis of a record in a database table, with the longitude and latitude of the center of each cell being stored as the reference location. GIS data maps for all of the variables noted above were all converted to raster format. The meteorological, geological, geographical, and demographic information for the center of each of the grid cells was extracted and added to the matching record in the database table. The estimated person-time at risk and the number of RVF cases occurring within each cell were added to the database records for each cell. The cell size of 3.5 km was chosen to match the spatial resolution of the majority of the data sources. Cell sizes for the 10-day NDVI and rainfall maps were 8 km. The centroids for these maps were overlaid on the 3.5-km cell map, and the 8-km centroid value closest to the 3.5-km centroid value was used for NDVI and rainfall. These maps, created for the East Africa horn region,¹⁹ were trimmed using the Kenya national border as a mask for these analyses. ArcView 9.3 was used for all geographical analyses.²²

Total person-time at risk for all people living within a grid cell was estimated in two steps: first, by estimating the population in the cell, and second, by multiplying this estimated population by the time that it was considered to be at

risk. The population of the grid cell was estimated by multiplying the population density of the sublocation in which the centroid of grid cell resided by the area of the grid cell (12.25 km^2).

Case locations were geocoded (assigned to specific grid cells) based on the location of the village or the centroid of the sublocation, location, or division in which the village was located if the location of the village could not be located. An iterative procedure similar to the procedure described in the work by Anyamba and others²³ was used to geocode case locations from three online gazetteers.

Locations for 214 cases were available and assigned global positioning system (GPS) locations (63% of 340 confirmed and probable cases). Reasons for cases not being geocoded were that either the village name could not be found in databases of village names or the case did not have a village, sublocation, location, or division reported.

Statistical methods. Statistical models were produced using Poisson regression. The number of cases ($N = 214$) in each 3.5-km-sided grid cell was the outcome variable, and geographic, geologic, and meteorological data for that cell were the explanatory variables. Offset values were the natural logarithm of the estimated person-time at risk for the cell defined as the estimated population in each 3.5-km cell multiplied by the time at risk (approximately 4 months).

Estimates from these models can be interpreted as relative risks. Combined with population information and risk factors for each grid cell, the model can be used to produce estimated incidences for each grid cell during the epidemic period. Model-based estimates of incidence per million person-years were obtained for each grid cell for the period at risk (Figure 1). Figure 1 insets show the Baringo and Kilifi districts. Backwards elimination was used to select significant variables in the multivariable model. However, because of collinearity with soil types, factors that were subcharacteristics of soil types were assessed together in models without soil types as a predictor. Fits of these models were compared using the Akaike Information Criterion.²⁴ The Poisson regression model estimated a relative risk per 0.1 unit change in the NDVI and per centimeter change in rainfall.

All models were examined for overdispersion; however, none existed. SAS V9.2 was used for statistical analyses.²⁵

The χ^2 test of independence (uncorrected) was used to compare case ($N = 70$) and non-case cells ($N = 46,130$) categorized by geologic and geographic variables. These bivariable results are comparable with the analytic methods used for the soil type results presented previously.⁴ However, because the population is unevenly spread across the country with respect to these analysis variables, there are differences between the location- and incidence-based analyses.

Table 1 presents the percent distribution of case and non-case locations with respect to potential explanatory variables, the number of cases, person-time during the outbreak period, and bivariable relative risks based on the levels of each variable used in the Poisson regression models.

RESULTS

The 214 cases occurred in 70 distinct grid cells; 171 cases in the North Eastern province occurred in 51 grid cells, 30 Baringo district cases occurred in 8 distinct grid cells, and 13 Kilifi district cases occurred in 11 distinct grid cells. Case

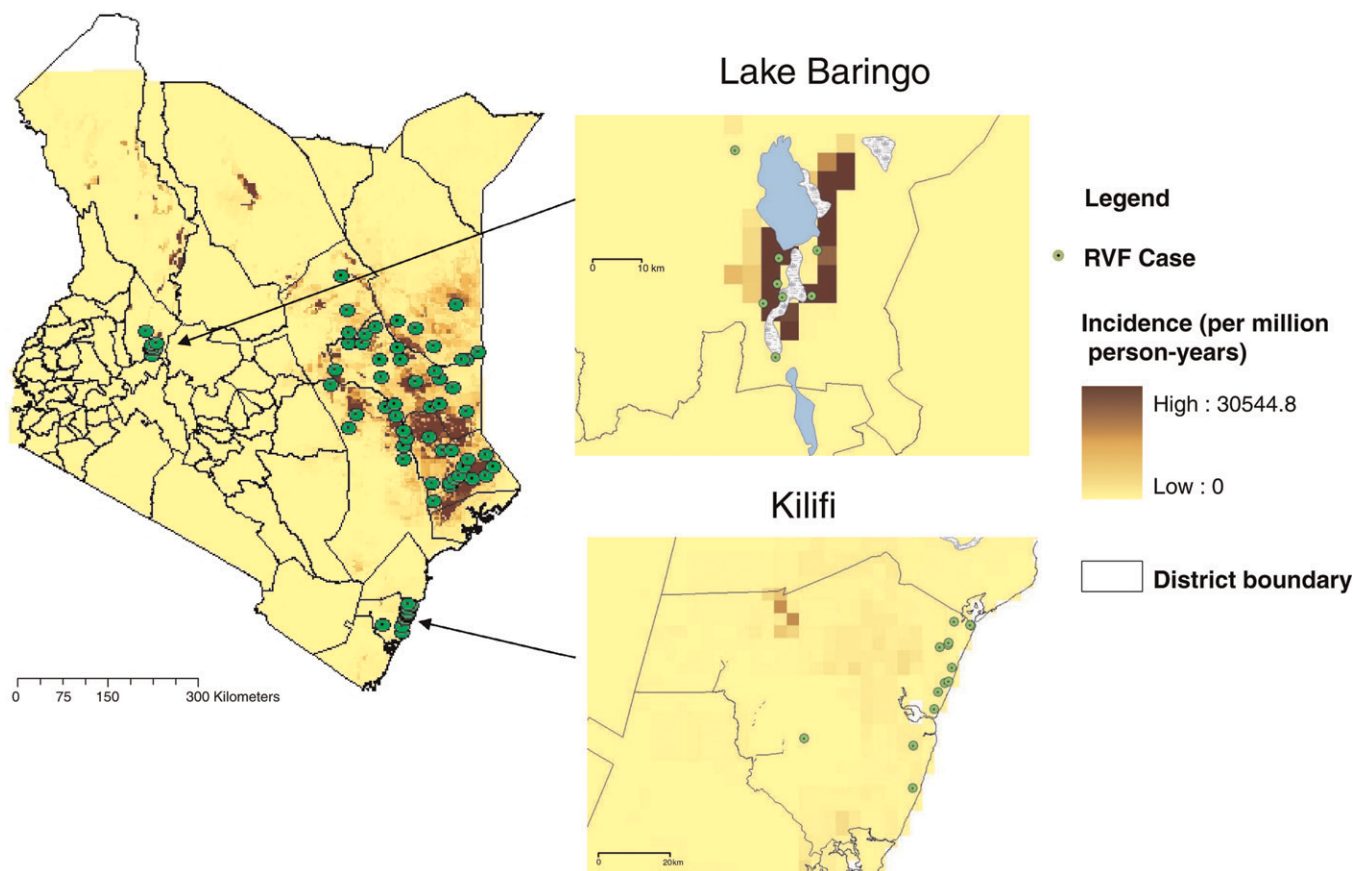


FIGURE 1. Estimated RVF incidence based on the multivariable model (November 1, 2006 to February 28, 2007).

locations are shown on the maps in Figure 1. Cases can be seen to be occurring in three clusters: (1) the North Eastern province in the eastern part of the cluster (this cluster is the spatially broadest cluster), (2) the Kilifi district cluster on the southeastern coast, and (3) the Baringo district cluster in the west-central part of the country.

Case and non-case grid cells were not significantly different with respect to populations or population densities (relative risk [RR] for population density = 0.9993; 95% confidence interval [CI] = 0.9977, 1.0003; $P = 0.30$).

BIVARIABLE RESULTS

Soil type. Four types of soils (solonetz, calcisols, solonchaks, and planosols) had statistically significantly elevated RRs versus all other soil types (Table 1). The soil types, using the three-letter subtypes from the FAO soil taxonomy,¹⁶ that were associated with increased RVF incidence were solonetz, calcisols, solonchaks, and planosols. Solonchaks, planosols, solonetz, and calcisols were present in areas where in 72% of the grid cells, cases occurred versus in 28% of the grid cells, no cases occurred ($P < 0.0001$, χ^2 test).

Soil texture. Soil types were classified by their texture properties (very clayey, clay, loamy, and sandy). These types were, in turn, recoded into clay soils versus all other soils. RVF incidence was higher in areas with clay or very clayey soils (Table 1). Clay soils were present in areas were 93% of the

cases occurred (data not shown) versus 70% for non-case locations ($P < 0.0001$, χ^2 test).

Soil drainage. Soil types were classified by their drainage properties and classified into three categories for the purposes of this analysis: extremely slow, very slow, and all other (which combines drainage group categories slow, well, rapid, and very rapid). The rapid and very rapid soil drainage groups are uncommon in Kenya. Areas having soils with both extremely slow and very slow drainage had increased risk of having RVF cases (Table 1).

Landforms. RVF incidence was higher in landforms categorized as plains (Table 1). Plains were 93% of case locations versus 74% of non-case cells ($P < 0.0001$, χ^2 test).

Ecological zones. RVF incidence was higher in grid cells categorized as Somali acacia (Table 1). These areas, primarily located in the North Eastern province, are semiarid areas where shrubby grasses and occasionally, trees are growing.

Land use. Land usage was classified as agricultural (dense and sparse), barren, forest, grass lands, bushlands (dense and sparse), and other. RVF case grid locations were more likely to be sparse agriculture or dense shrubby land use locations versus non-case locations (73% versus 58%; $P < 0.0001$, χ^2 test).

Elevation. The majority of the cases occurred at elevations below 500 m (88.6% for case locations versus 36.0% for non-case locations) (Table 1). No case occurred above 1,100 m. Approximately 30% of Kenya is at an elevation of 1,000 m or higher.

TABLE 1
Geographic and geologic factors

Variable	Case locations percent	Non-case locations percent	Cases (percent)	Person-time years	Incidence per million person-years	Relative risk	Lower 95% CL	Upper 95% CL
Soil type	<i>N</i> = 70	<i>N</i> = 46,130	<i>N</i> = 214	<i>N</i> = 8,582,024				
Solonetz	34.3	12.4	74 (34.6)	120,547	613.9	105.8	73.5	154.3
Calcisols	4.3	10.4	11 (5.1)	76,152	144.8	24.5	12.1	45.7
Solonchak	7.1	2.3	25 (11.7)	34,433	726.0	127.1	76.9	205.4
Planosols	15.7	5.6	57 (26.6)	259,472	219.7	39.3	26.6	58.3
All other	38.6	69.4	47 (22.0)	8,091,419	5.8	Reference		
Soil texture	<i>N</i> = 70	<i>N</i> = 45,868	<i>N</i> = 214	<i>N</i> = 8,582,024				
Clay	87.1	67.0	197 (92.1)	6,378,268	30.9	3.8	2.4	6.6
All other	12.9	33.0	17 (7.9)	2,145,311	7.9	Reference		
Soil drainage	<i>N</i> = 70	<i>N</i> = 45,868	<i>N</i> = 214	<i>N</i> = 833,579				
Extremely slow	21.4	27.6	75 (35.0)	1,875,380	40.0	3.5	2.2	5.9
Very slow	58.6	50.1	117 (54.7)	4,896,064	23.90	2.1	1.3	3.5
All other	20	22.3	22 (10.3)	1,762,134		Reference		
Landform	<i>N</i> = 70	<i>N</i> = 46,130	<i>N</i> = 214	<i>N</i> = 8,582,024				
Plains	92.9	74.0	203 (94.9)	4,472,875	45.4	16.7	9.6	36.5
All other	7.1	25.98	11 (5.1)	4,109,149	2.68	Reference		
Ecozone	<i>N</i> = 70	<i>N</i> = 46,130	<i>N</i> = 214	<i>N</i> = 8,582,024				
Somali acacia	10.0	16.6	22 (10.3)	88,240	249.3	10.3	6.4	15.7
All other	90.0	83.4	192 (89.7)	8,493,784	22.6	Reference		
Land use	<i>N</i> = 70	<i>N</i> = 46,003	<i>N</i> = 214	<i>N</i> = 8,553,042				
Sparse agriculture	15.7	17.4	31 (14.5)	273,298	113.4	20.0	12.5	31.8
Dense bush	57.1	41.9	140 (65.4)	881,093	158.9	28.2	20.1	40.5
All other	27.14	40.7	43 (20.1)	7,398,651	5.8	Reference		
Elevation (m)	<i>N</i> = 70	<i>N</i> = 45,820	<i>N</i> = 214	<i>N</i> = 8,559,500				
0–500	88.6	36.0	184 (86.0)	816,663	225.3	259.0	125.7	658.1
501–1,000	7.1	34.5	24 (11.2)	646,245	37.1	43.0	18.74	116.0
1,000+	4.3	29.5	6 (2.8)	7,096,593	0.85	Reference		

CL = confidence limit.

Rainfall. Locations that had cases of RVF had significantly less annual rainfall than non-case locations (49.1 versus 57.5 cm/year; RR per cm = 0.947; 95% CI = 0.942, 0.952). On a purely temporal basis, case locations had significantly greater rainfall than non-case locations for the periods of November 1–10, November 11–20, and December 11–20 (all in 2006). Case locations had significantly less rainfall than non-case locations for the periods of November 21–30 and December 1–10 and from December 21 to February 28. Associations for rainfall for 10-day periods by location are presented in a subsequent section. For the period of February 11–20, there was no rainfall in any case location, and therefore, the RR was undefined. However, a χ^2 test for any versus no rainfall for case and non-case locations was highly significant ($P < 0.0001$) for this time period. Rainfall amounts were over 5 cm for case locations for each 10-day period in November but dropped dramatically after January 1, 2007, only once exceeding 0.5 cm for any period. Bivariable results are presented in Table 2. Maps of rainfall in Kenya for selected 10-day periods are shown in Figure 2.

NDVI. Case locations had significantly higher NDVIs than non-case locations during the periods of November 11–20, November 21–30, and December 21–31. Non-case locations had significantly higher NDVIs during the periods of November 1–10, December 1–10, and December 11–20 and from January 1, 2007 to February 28, 2007 (Table 2). Greenness measures for case locations declined dramatically after January 1, 2007.

Multivariable model. Soil types were highly significant in the multivariable model (Table 3). Solonchaks, as found near Lake Baringo, had the highest RR (RR = 96.4; 95%

CI = 48.1, 195.7) of any other factor in the model. Most rainfall measures remained statistically significant. As a group, the rainfall and soil types were the most significant factors in the model. A number of the NDVI measures became non-significant and were dropped from the model without any impact on the estimates of the remaining variables in the model. Also remaining significant in the multivariable model were plains areas, elevation, densely bushy areas, and the Somalia acacia ecozone, which occurs primarily in the North Eastern province.

Incidence. Locations that had baseline values for all variables in the multivariable model, including zero values for elevation, rain, and NDVI, could be expected to have an incidence of 10.8 cases per million person-years (95% CI = 4.3, 25.2). Figure 1 presents model-based incidences per million person-years based on the attributes of the grid cell.

Findings by location. Because cases occurred in spatial and temporal clusters in North Eastern province, Kilifi district, and Baringo district, bivariable analyses, where the cases from each cluster were compared with all other non-case location, were performed. Analyses were limited by the small numbers of cases. Baringo case location risk factors were dense bush land use, solonchak soil type, and rain in the first 10 days of February. Rainfall in Baringo and western Kenya was much higher than other parts of the country, including other case locations during this time period (Baringo case locations: 4.9 ± 0.6 cm; other case locations: 0.01 ± 0.05 cm; non-case locations: 1.3 ± 2.1 cm) (Figure 2). Rain during the first 10 days of November was the only significant risk factor for Kilifi locations. Risk factors for North Eastern province case locations included rainfall during November 1–10, November 11–20,

TABLE 2
Rainfall and NDVI measures by month and dekads (10-day interval)

Rainfall (cm)	Case locations			Non-case locations			Relative risk	Lower 95% CL	Upper 95% CL
	<i>N</i>	Mean	SD	<i>N</i>	Mean	SD			
Rain (annual)	69	49.1	24.6	45,813	57.5	31.8	0.95	0.94	0.95
November 1–10	70	6.74	5.56	46,130	3.66	3.50	1.11	1.08	1.14
November 11–20	67	10.06	4.42	45,532	6.36	5.00	1.06	1.03	1.08
November 21–30	67	5.56	1.51	45,532	6.37	3.27	0.74	0.70	0.78
December 1–10	67	1.44	1.94	45,532	2.56	2.38	0.71	0.66	0.78
December 11–20	67	7.09	5.83	45,532	3.15	4.10	1.26	1.23	1.30
December 21–31	67	2.14	2.63	45,532	5.46	5.39	0.66	0.63	0.69
January 1–10	67	0.41	0.61	45,532	1.34	1.80	0.33	0.27	0.40
January 11–20	67	0.43	0.50	45,532	0.79	1.72	0.51	0.44	0.58
January 21–31	67	0.25	0.68	45,532	0.59	1.07	0.32	0.25	0.40
February 1–10	67	0.59	1.60	45,532	1.31	2.12	0.65	0.59	0.70
February 11–20	67	0	0	45,532	0.27	0.67	Undefined	Undefined	Undefined
February 21–28	67	0.48	0.52	45,532	0.99	1.15	0.15	0.11	0.19
NDVI									
November 1–10	67	−0.16	0.26	45,532	−0.22	0.37	0.996	0.959	1.034
November 11–20	67	−0.06	0.24	45,532	−0.12	0.38	1.023	0.980	1.07
November 21–30	67	0.020	0.29	45,532	0.01	0.47	1.0425	1.009	1.076
December 1–10	67	0.10	0.29	45,532	0.07	0.41	0.965	0.921	1.015
December 11–20	67	−0.12	0.31	45,532	−0.02	0.39	0.849	0.818	0.883
December 21–31	67	0.11	0.35	45,532	0.09	0.46	1.1345	1.098	1.172
January 1–10	67	0.12	0.31	45,532	0.12	0.37	0.922	0.879	0.971
January 11–20	67	−0.09	0.38	45,532	−0.08	0.40	0.895	0.861	0.932
January 21–31	67	−0.16	0.28	45,532	−0.080	0.361	0.792	0.765	0.820
February 1–10	67	−0.25	0.26	45,532	−0.183	0.399	0.780	0.754	0.808
February 11–20	67	−0.21	0.23	45,532	−0.192	0.362	0.789	0.761	0.820
February 21–28	67	−0.96	0.04	45,532	−0.923	0.090	0.086	0.058	0.122

and December 11–20, NDVI of December 21–31, plains landforms, and solonetz, planosol, and calcisol soil types.

DISCUSSION

The association of a number of past RVF outbreaks with flooding has led to a model to forecast future outbreaks.^{11,12,23} Notably, the 2006–2007 RVF outbreak in east Africa was forecast by one of these models.²³ For our model, geocoded case locations, geographic and other geocoded information for the entire country, census data, and a defined period of risk allows Poisson regression to be used for modeling, which in turn, allows RRs and incidences to be computed. It should be noted that our model uses variables to predict RVF incidence during the outbreak period rather than outbreaks of RVF.

Taken as a whole, the variables in the final multivariable model associate increased incidence of RVF with locations that have attributes that provide optimal vector habitat at each life stage. Elevation reflects the limited range of the vectors. The lower NDVI in the first 10 days of November of 2006 describes an area that is more arid than the rest of the country. The increased rainfall preceding the outbreak period provides water to rehydrate desiccated mosquito eggs in soil. A dense bush vegetation cover could provide landing zones and resting areas that would be desirable to vectors. The plains landform allows flood waters to pool more easily to provide larval habitat than hilly or other non-flat landforms. The association of soil types has been discussed in detail previously, including with a soil map of Kenya showing case locations.⁴ Briefly, all of the associated soil types have substrata that could serve to retain water better than other soil types (e.g., sandy), a feature that could plausibly facilitate rehydration of desiccated

mosquito eggs in normally arid settings. When investigating relationships between RVF and soil types in other settings, consideration should be given to any soil type that forms water-retaining strata, not just the types that were found to be associated in this outbreak investigation.

Our model provides both linkages and contrasts between the cases occurring in the three clusters. The soil type analysis provides a potential linkage between the Baringo district cluster, where cases occurred in solonchak soils in a wetlands, to the cluster in normally arid North Eastern province, where solonetz soils, among others, were associated with case occurrence. Solonchak soils transition to solonetz soils on drying.^{4,16} In our models for each cluster, the North Eastern and Kilifi cases were both associated with rainfall occurring approximately 3 weeks before the first reported case onset. The first case in North Eastern province occurred on November 30, 2006, whereas the first Kilifi case occurred on December 1, 2006.⁴ The Baringo model showed that cases were associated with increased rainfall in early February when the rest of the country was much drier (including other case locations). The first Baringo case occurred on January 25, with peak occurrence during the first week of February⁴ (Figure 2).

Both rainfall and NDVI measures were in our final model, and their roles are complimentary. The coefficients of NDVI and the rainfall amounts for the first 10 days of November describe an arid area receiving more rainfall than other parts of the country. Although they are related, both factors are important in different ways in the mosquito lifecycle. Rainfall would be important in rehydrating soils needed to help mosquito eggs hatch, whereas higher values of NDVIs (with appropriate land cover) could reflect better resting places for mosquitoes. Higher values of NDVIs are often most correlated with rainfall occurrence in preceding weeks,

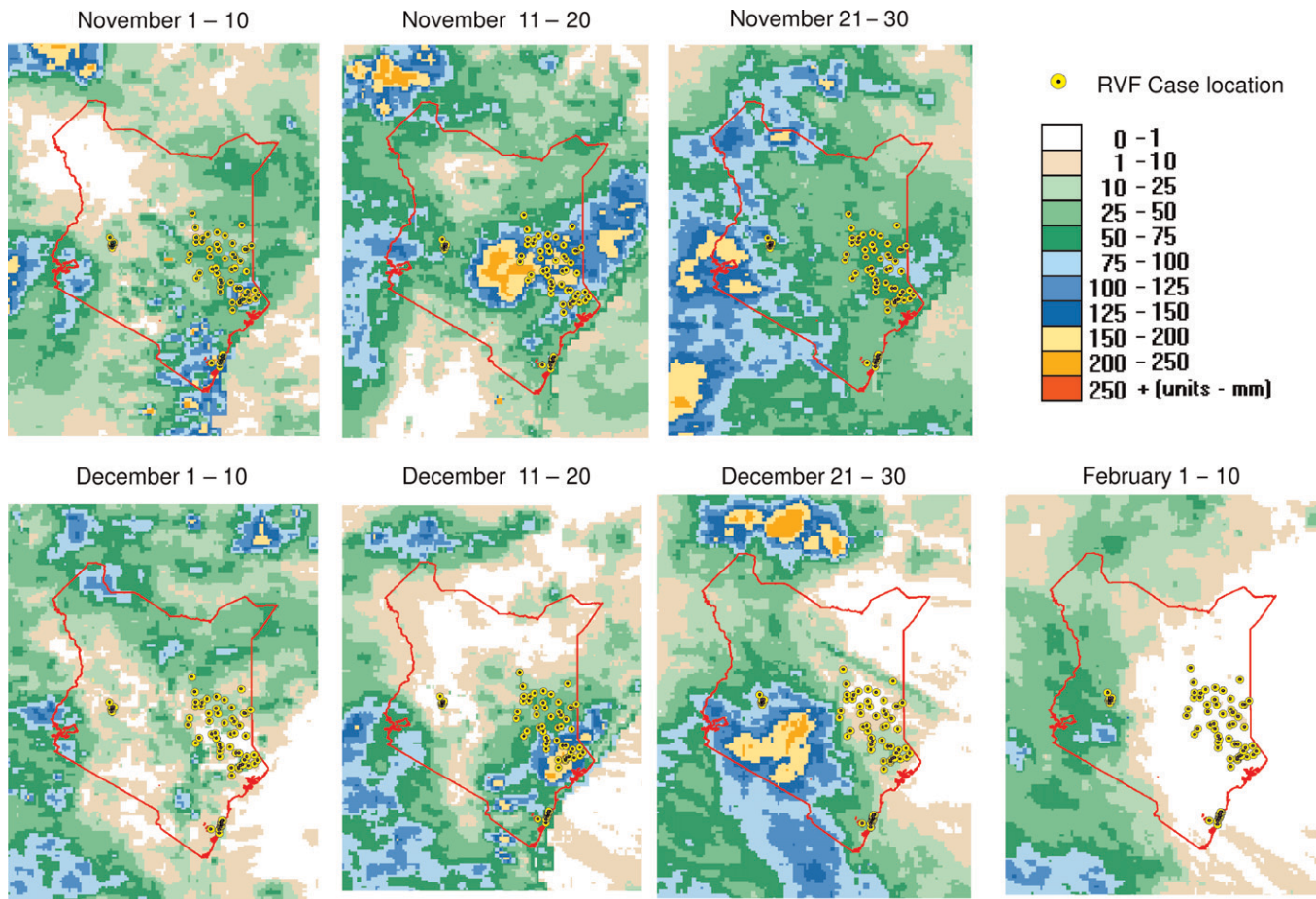


FIGURE 2. Rainfall in selected 10-day periods in Kenya. The first epidemic period, involving North Eastern province and Kilifi district, occurred from mid-December of 2006 to mid-January of 2007. The second epidemic period, involving Baringo district, began in mid-January of 2007 and peaked in early February.

particularly during the beginning of rainy seasons. In arid areas, increases in rainfall precede increases in vegetation cover.¹⁹ As a result, increases in rainfall might be a better early indicator than NDVI measures for outbreak prediction.

The estimated incidences in Figure 1 suggest that most of the country was at low risk for RVF during the outbreak period. There are some areas of estimated high incidence in northwestern and northeastern Kenya, where no cases were reported during this outbreak. These areas generally have solonetz (in the northeast) or solonchak (in the northwest) soil types,⁴ are plains, are sparsely populated, and in the case of the northeastern risk areas, are in the Somalia acacia ecological zone. The model notably does not explain the Kilifi cases well. Rainfall in early November was the only significant predictor of case occurrence in this area among the variables considered. There are known informal trade routes for livestock originating in North Eastern province that pass through coastal areas, including Kilifi district.²⁶ If some of these infected animals were the basis of the human disease, then livestock would provide a simple explanation for why these cases were not explained by climate or geology.

This modeling approach has several shortcomings. First, it does not account for host susceptibility levels in both the human and animal populations; animals and humans previously exposed to RVF virus would be unlikely to contribute

to propagating and spread of virus during a period of potential virus transmission. This result could overpredict outbreak occurrence or incorrectly assess an individual's risk. Second, cases could have occurred in areas with no surveillance or reporting capabilities. Omission of such unreported cases could have easily changed the findings of the model. Third, it does not include individual risk activities, such as contact with bodily fluids from infected animals.^{3,9} Geographic locations for cases were approximate and may not be the same as their location when infected. This finding could result in attributes for case locations being biased to those attributes for non-case locations. Finally, this model used population data based on an assumption of a uniform density of people within the smallest administrative unit, the sublocation. Severe violations of this assumption could lead to inaccuracies in the estimates of incidence.

Analyses that were based on case versus non-case locations alone yielded somewhat different results than analyses using person-time as denominators (Table 1). This difference is because of the distribution of the population being different from the distribution of geological factors. For example, calcisols were a smaller percent of case locations than non-case locations, making it seem to be a risk factor for the location-based analysis. However, few people lived in areas with calcisols (less than 1% of person-time, despite covering

TABLE 3
Variables in the multivariable model predicting RVF incidence

Parameter	Relative risk	Likelihood ratio 95% confidence limits		Wald χ^2	Pr > χ^2
		Lower	Upper		
Intercept	1.08E-05	4.3E-06	2.5E-05	656.47	< 0.0001
Soil type					
Solonetz	3.26	2.02	5.37	22.49	< 0.0001
Calcisols	4.17	1.85	8.78	13.09	0.0003
Solonchaks	96.44	48.14	195.72	163.8	< 0.0001
Planosols	10.70	6.49	17.96	83.56	< 0.0001
Other	Reference				
Land use					
Dense bush	1.79	1.28	2.53	11.29	0.0008
Other	Reference				
Landform					
Plains	2.04	1.10	4.14	4.49	0.0341
Other	Reference				
Ecological zone					
Somali acacia	3.88	2.17	6.75	22.18	< 0.0001
Other	Reference				
Rain* (cm)					
November 1–10	1.08	1.05	1.12	22.54	< 0.0001
November 11–20	1.24	1.19	1.30	82.69	< 0.0001
November 21–30	0.72	0.65	0.80	40.45	< 0.0001
December 11–20	1.18	1.13	1.23	50.74	< 0.0001
January 1–10	0.37	0.27	0.48	48.97	< 0.0001
February 1–10	2.42	1.88	3.17	43.74	< 0.0001
NDVI†					
November 1–10	0.83	0.72	0.96	6.73	0.0095
November 11–20	1.25	1.08	1.43	9.14	0.0025
January 11–20	1.09	1.04	1.14	16.24	< 0.0001
Elevation per 100 m	0.68	0.59	0.76	36.13	< 0.0001

* Per centimeter change.

† Per 0.1 unit change.

10% of Kenya), resulting in the incidence for those areas being significantly greater than the reference areas. Similarly, the differences in the percent of case and non-case locations that were solonetz soil types were relatively small. As with calcisols, few people lived in these areas, which when combined with the number of cases occurring, resulted in a high incidence and increased RRs during the outbreak period.

One of the strengths of this model is that it uses geographic information that is commonly available. Many countries have rich geographical, geological, and meteorological data available in georeferenced format that could serve as a foundation for similar investigations of the occurrence of RVF or other diseases. Although considerable effort was spent geocoding the case locations, the specificity of this information allowed optimal use of the available reference data on the other geographic variables. This effort is the first that takes place on a scale between individual risk factors^{3,4,9} and factors affecting multiple countries or regions.²³

It improves on earlier efforts⁴ by using methods that allow simultaneous assessment of multiple variables and incidences for the epidemic period to be computed.

In conclusion, our models suggest that RVF incidence during the outbreak period in Kenya was related to a number of geological, geographical, and meteorological factors, some previously recognized and others not recognized. We have shown that the Kilifi and North Eastern clusters could be linked by rainfall in early November and that the Baringo and North Eastern clusters could be linked by soil types and land cover. The model notes that all cases occurred in generally flat areas and at lower altitudes (always 1,100 m or less).

The findings suggest that, although rainfall and associated measures are important predictors of RVF outbreaks, there are additional factors that better define the optimal environment for RVF occurrence. Such findings have the potential to improve current outbreak prediction models by limiting the geographic ranges of prediction to areas that are at risk of having outbreaks occur.

Received July 12, 2011. Accepted for publication November 21, 2011.

Acknowledgments: The authors would like to thank Dr. Solomon Mpoke, Director of the Kenya Medical Research Institute, John Vulule, Director of the Centre for Global Health Research, Kenya Medical Research Institute and the GIS Users Group of the Centers for Disease Control and Prevention for their support of this work.

Disclaimer: The findings and conclusions in this report are the authors and do not necessarily represent the views of the Centers for Disease Control and Prevention.

Authors' addresses: Allen Hightower (retired), Division of Parasitic Diseases and Malaria, Centers for Disease Control and Prevention, Atlanta, GA, E-mail: awh1953@gmail.com. Carl Kinkade, Epidemiology and Analysis Program Office, Office of Surveillance Epidemiology and Laboratory Systems, Centers for Disease Control and Prevention, Atlanta, GA, E-mail: mke5@cdc.gov. Patrick M. Nguku and David Mutonga, Division of Disease Surveillance and Response, Ministry of Public Health and Sanitation, Nairobi, Kenya, E-mails: drnguku@yahoo.com and davidmutonga@yahoo.com. Amwayi Anyangu, Field Epidemiology and Laboratory Training Program (FELTP), Department of Disease Prevention and Control, Ministry of Public Health and Sanitation, Nairobi, Kenya, E-mail: amwayi2004@yahoo.com. Jared Omolo, Field Epidemiology and Training Program, Ministry of Public Health and Sanitation, Nairobi, Kenya, E-mail: jaredom2000@yahoo.com. M. Kariuki Njenga, International Emerging Infections Program, Centers for Disease Control and Prevention–Kenya, Kenya Medical

Research Institute Headquarters, Nairobi, Kenya, E-mail: KNjenga@ke.cdc.gov. Daniel R. Feikin, Johns Hopkins Bloomberg School of Public Health, Baltimore, MD, E-mail: dfeikin@jhsph.edu. David Schnabel, US Army Medical Research Unit–Kenya, Nairobi, Kenya, E-mail: david.c.schnabel@us.army.mil. Maurice Ombok, Center for Global Health Research, Kenya Medical Research Institute, Kisumu, Kenya, E-mail: mombok@ke.cdc.gov. Robert F. Breiman, Centers for Disease Control and Prevention–Kenya, Kenya Medical Research Institute Headquarters, Nairobi, Kenya, E-mail: RBreiman@ke.cdc.gov.

REFERENCES

- Peters CJ, Linthicum KJ, 1994. Rift Valley fever. Beran GW, ed. *CRC Handbook Series in Zoonoses, Section B: Viral Zoonoses*, 2nd ed. Boca Raton, FL: CRC Press Inc., 125–138.
- Davies FG, Linthicum KJ, James AD, 1985. Rainfall and epizootic Rift Valley fever. *Bull World Health Organ* 63: 941–943.
- Woods CW, Karpati AM, Grein T, McCarthy N, Gaturuku P, Muchiri E, Dunster L, Henderson A, Khan Ali S, Swanepoel R, Bonmarin I, Martin L, Mann P, Smoak BL, Ryan M, Ksiazek TG, Arthur RR, Ndikuyeze A, Agata NN, Peters CJ; World Health Organization Hemorrhagic Fever Task Force, 2002. An outbreak of Rift Valley Fever in northeastern Kenya, 1997–98. *Emerg Infect Dis* 8: 138–144.
- Nguku PM, Sharif SK, Mutonga D, Amwayi S, Omolo J, Mohammed O, Farnon EC, Gould LH, Lederman E, Rao C, Sang R, Schnabel D, Feikin DR, Hightower A, Njenga MK, Breiman RF, 2010. An investigation of a major outbreak of Rift Valley Fever in Kenya: 2006–2007. *Am J Trop Med Hyg* 83 (Suppl 2): 5–13.
- Linthicum KJ, Davies FG, Kairo A, Bailey CL, 1985. Rift Valley fever virus (family Bunyaviridae, genus *Phlebovirus*): isolations from Diptera collected during an interepizootic period in Kenya. *J Hyg (Lond)* 95: 197–209.
- Logan TM, Linthicum KJ, Davies FG, Binopal YS, Roberts CR, 1991. Isolation of Rift Valley fever virus from mosquitoes collected during an outbreak in domestic animals in Kenya. *J Med Entomol* 28: 293–298.
- Meegan JM, Khalil GM, Hoogstrahl H, Adham FK, 1980. Experimental transmission and field isolation studies implicating *Culex pipiens* as a vector of Rift Valley fever virus in Egypt. *Am J Trop Med Hyg* 29: 1405–1409.
- Sang R, Kioko E, Lutomia J, Warigia M, Ochieng C, O’Guinn M, Lee JS, Koka H, Godsey M, Hoel D, Hanafi H, Miller B, Schnabel D, Breiman RF, Richardson J, 2010. Rift Valley Fever virus epidemic in Kenya, 2006/2007: the entomologic investigations. *Am J Trop Med Hyg* 83: 28–37.
- Anyangu A, Gould LH, Sharif SK, Nguku PM, Omolo JO, Mutonga D, Rao CY, Lederman ER, Schnabel D, Paweska JT, Katz M, Hightower A, Njenga MK, Feikin DR, Breiman RF, 2010. Risk factors for severe Rift Valley Fever infection in Kenya, 2007. *Am J Trop Med Hyg* 83: 14–21.
- Munyua P, Murithi R, Wainwright S, Githinji J, Hightower A, Mutonga D, Macharia J, Ithondeka P, Musaa J, Breiman R, Bloland P, Njenga MK, 2010. Rift Valley Fever outbreak in livestock in Kenya, 2006–2007. *Am J Trop Med Hyg* 83: 58–64.
- Linthicum KJ, Anyamba A, Tucker CJ, Kelley PW, Myers MF, Peters CJ, 1999. Climate and satellite indicators to forecast Rift Valley fever epidemics in Kenya. *Science* 285: 397–400.
- Anyamba A, Chretien J, Small J, Tucker CJ, Formenty P, Richardson JH, Britch SC, Schnabel DC, Erickson RL, Linthicum KJ, 2009. Prediction of a Rift Valley fever outbreak. *Proc Natl Acad Sci USA* 106: 955–959.
- Clements A, Pfeiffer D, Martin V, Thiongane Y, Sall B, Pittiglio C, 2007. Risk-mapping of Rift Valley fever in Senegal. *Vector Borne Zoonotic Dis* 7: 203–216.
- Bicout DJ, Sabatier P, 2004. Mapping Rift Valley fever vectors and prevalence using rainfall variations. *Vector Borne Zoonotic Dis* 4: 33–42.
- Kruska A, Kariuki P, 2007. International Livestock Research Institute GIS Database. Available at: <http://www.ilri.org/gis>. Accessed December 2, 2010.
- FAO, 1988. *FAO/UNESCO Soil Map of the World, Revised Legend with Corrections and Updates. World Soil Resources Report 60*. Rome, Italy: FAO, Rome.
- Olson DM, Dinerstein E, Wikramanayake ED, Burgess ND, Powell GVN, Underwood EC, D’Amico JA, Itoua I, Strand HE, Morrison JC, Loucks CJ, Allnutt TF, Ricketts TH, Kura Y, Lamoreux JF, Wettengel WW, Hedao P, Kassem KR, 2001. Terrestrial ecoregions of the world: a new map of life on Earth. *Bioscience* 51: 933–938.
- World Resources Institute, Ministry of Environment and Natural Resources, International Livestock Research Institute, Kenya National Bureau of Statistics, 2007. Nature’s Benefits in Kenya. Available at: <http://www.wri.org/publication/content/9291>. Accessed December 2, 2010.
- Nicholson SE, Davenport ML, Malo AR, 1990. A comparison of the vegetation response to rainfall in the Sahel and East Africa, using normalized difference vegetation index from NOASS AVHRR. *Clim Change* 17: 209–241.
- Africa Data Dissemination Service, 2007. FEWS NET Data Portal. Available at: <http://earlywarning.usgs.gov/fews/africa/index.php>. Accessed December 2, 2010.
- Kenya Central Bureau of Statistics, 2001. The 1999 Population & Housing Census: Population by Administrative Areas and Urban Centres. Available at: <http://www.knbs.or.ke/>. Accessed December 2, 2010.
- Environmental Systems Research Institute, Inc., 2008. *ArcView Version 9.3 Software*. Redlands, CA: ESRI.
- Anyamba A, Linthicum KJ, Small J, Britch SC, Pak E, de La Rocque S, Formenty P, Hightower AW, Breiman RF, Chretien J, Tucker CJ, Schnabel D, Sang R, Haagsma K, Latham M, Lewandowski HB, Magdi SO, Mohamed MA, Nguku PM, Reynes J, Swanepoel R, 2010. Prediction, assessment of the Rift Valley Fever activity in East and Southern Africa 2006–2008 and possible vector control strategies. *Am J Trop Med Hyg* 83: 43–51.
- Bozdogan H, 1987. Model selection and Akaike’s Information Criterion (AIC): the general theory and its analytical extensions. *Psychometrika* 42: 345–370.
- SAS, 2008. *The Statistical Analysis System (SAS) Version 9.2 Software*. Cary, NC: SAS.
- Muthee A, 2006. Kenya Livestock Sector Study. Available at: <http://www.mmw4p.org/dyn/bds/docs/552/KenyaLivestockValueChainReport.pdf>. Accessed September 23, 2011.