Identification of case clusters and counties with high infective connectivity in the 2001 epidemic of foot-and-mouth disease in Uruguay

Gerardo Chowell, PhD; Ariel L. Rivas, DVM, PhD; Stephen D. Smith, MS; James M. Hyman, PhD

Objective—To evaluate the influence of individual spatial units (ie, counties) on the epidemic spread of foot-and-mouth disease (FMD) virus.

Sample Population—163 counties in Uruguay where there was an outbreak of FMD between April 23 and July 11, 2001.

Procedure—A geographically referenced database was created, and the distance between counties (13,203 county pairs), road density of counties (163 counties), and time when cases were reported in those counties (11 weeks of the epidemic) were considered to assess global spatial and spatial-temporal autocorrelation, determine the contribution of links connecting pairs of counties with infected animals, and allow us to hypothesize the influence for spread during the epidemic for counties with greater than the mean infective link contributions.

Results—Case clusters were indicated by the Moran *I* and Mantel tests during the first 6 weeks of the epidemic. Spatial lags between pairs of counties with infected animals revealed case clustering before and after vaccination was implemented. Temporal lags predicted autocorrelation for up to 3 weeks. Link indices identified counties expected to facilitate epidemic spread. If control measures had been implemented in counties with a high index link (identifiable as early as week 1 of the epidemic), they could have prevented (by week 11 of the epidemic) at least 2.5 times as many cases per square kilometer than the same measures implemented in counties with average link indices.

Conclusions and Clinical Relevance—Analysis of spatial autocorrelation and infective link indices may identify network conditions that facilitate (or prevent) disease spread. (*Am J Vet Res* 2006;67:102–113)

The emergence of GISs has opened a new method for analyzing spatial dynamics of epidemics.[1] Spatial features (ie, mountains, cities, rivers, and farms) are rarely distributed in random or regular patterns. They are usually fragmented (discontinuous). Spread of disease during an epidemic may be influenced by factors that include but go beyond topographic features (such as winds, human traffic, road density, and other spatial variables).[2,3]

An epidemic process may be regarded as composed of 2 spatial points (eg, 2 animals, 2 farms, or 2 counties) connected through a line. One of these points is the infector and the other the infected. The line may have multiple forms (eg, a road or a delivery route). By expanding this concept to that of a network (a set of nodes or points linked by multiple lines), animals located at nodes are expected to be infected during an epidemic that spreads along the lines. Hence, the issue of interest is to identify the unknown lines of an epidemic network.[4-7]

Spatial connectivity depends on Euclidean (straight line) and non-Euclidean distances (eg, connections through roads), which are factors that influence spread of disease during an epidemic.[8] Euclidean distance can be estimated by measuring the distance between centroids (eg, farm or county centroids).[9] Non-Euclidean distance can be assessed by estimating total (major and minor) road density, which tends to be linearly predicted by major road density.[10]

Epidemic spatial connectivity may be investigated by use of classic spatial statistical techniques. They include the Moran *I* test (which assesses spatial autocorrelation), Mantel test (which measures spatial-temporal autocorrelation), and their derived correlograms. The correlograms identify the distance or time lag within which spatial autocorrelations extend.[11,12] The Moran test evaluates whether there is a spatial autocorrelation (eg, whether cases are associated with sites spatially close to each other, such as in adjacent counties).[13] Positive autocorrelation exists when the magnitude of cases increases as spatial proximity increases. Similarly, the Mantel statistic is used to assess spatial and temporal autocorrelation.[14,15]

Although local Moran and Mantel tests can quantify the contribution of each specific spatial point to the overall (spatial or temporal-spatial) autocorrelation,[12] most local tests are not spatially explicit because they do not identify the line that connects an infected point to other (susceptible or subsequently infected) points. They are not spatially explicit or, if spatially explicit (ie, the scan statistic test), not appropriately suited to detect long-distance links (ie, not appropriate to detect fragmented clusters).[16-22] Those limitations could be addressed by local tests that focus on the connecting line between points. Connectivity has been

GIS Geographic information system
FMD Foot-and-mouth disease

Received March 14, 2005.
Accepted June 8, 2005.
From Mathematical Modeling and Analysis, Los Alamos National Laboratory, Los Alamos, NM 87545 (Chowell, Hyman); and the Department of Biological Statistics and Computational Biology (Rivas) and Institute for Resource Information Systems (Smith), College of Agriculture and Life Sciences, Cornell University, Ithaca, NY 14853.
Dr. Chowell was supported by a Director's Postdoctoral Fellowship from Los Alamos National Laboratory.
The authors thank Dr. Carlos Olave Castro and Capt. Norbertino Suárez for technical assistance.
Address correspondence to Dr. Rivas.

investigated from a network point of view (spatial link analysis) as conceptualized in a classic study and used in various fields.[4-7] Together, assessments of spatial-temporal autocorrelation, supplemented with local tests that estimate the contribution to the overall autocorrelation provided by specific connections (spatial links between pairs of infected locations), could spatially identify geographically proximal case clusters (close-distance connections) as well as "nonclustered" clusters (ie, cases that are located in spatially fragmented areas and connected by long-distance links).

To test the influence of spatial connectivity on disease dispersal during an epidemic, geographically referenced epidemic data are needed. The 2001 epidemic of FMD in Uruguay offers an opportunity to evaluate diffusion over time and space during an epidemic. Cattle were predominantly infected in a country previously free of FMD.[23-25] The minimal replication cycle of FMD virus is estimated to be 3 days.[26] Studies[27-29] on FMD and other diseases have indicated heterogeneous spatial spread and used the centroids of irregular polygons (ie, counties) as units of analysis. Road networks may influence dispersal of FMD virus.[24,25,30]

The study reported here had 3 objectives. We intended to determine whether infected sites were spatially or temporally autocorrelated; if sites were clustered, to measure the contribution of each spatial link to the overall spatial-temporal autocorrelation; and to use that information to generate and evaluate hypotheses on the various potentials for disease spread during an epidemic for specific counties.

Materials and Methods

Description of the epidemic—Details of this epidemic have been reported[23-25] elsewhere. Initial cases of FMD were identified in the southwestern quadrant of Uruguay, a nonurban, cattle-raising region characterized by higher road density than the national median (**Figures 1 and 2**). Several interventions were implemented over time, including a nationwide ban on animal movement (implemented on or after day 2 of the epidemic) and a nationwide program of vaccination. However, human traffic was not interrupted. Milk trucks continued to visit dairy farms and collect milk throughout the duration of the epidemic. In addition, no vaccines were available in the country at the time the epidemic began.[31,32] Although a decision to acquire > 10 million doses of vaccine was made within a week after the onset of the epidemic, no data were available in relation to where or when the first vaccination was implemented. It is estimated that at least 3 days are required for immunologically naïve animals to synthesize antibodies after vaccination with a high-potency vaccine.[33] No spatial-temporal data were available as to whether vaccine-induced antibodies reached protective titers. A second vaccination was implemented later.

GIS procedures—Two GIS packages[a,b] were used to geographically reference data and create maps. An official map of Uruguay,[c] including the location and area of the 276 counties, was used. On the basis of the 2000 Agricultural Census for Uruguay, 248 counties (cattle-raising regions) were selected. Of those, 163 counties contained infected animals at some time during the 11-week period that began on April 23, 2001. Geographically coded data on weekly (county level) and daily (for the first 6 days only; farm level) number of cases were retrieved from public sources and processed as described elsewhere.[24,34-37]

Four steps were used to determine the intercounty centroid distance. First, the x- and y-coordinates for each county's surface were identified by accessing the x- and y-values in the shape field. Second, the center value for each polygon (centroid) was provided by use of the GIS packages. Third, a point layer was generated from the x- and y-values of the centroid for each county. Fourth, distances

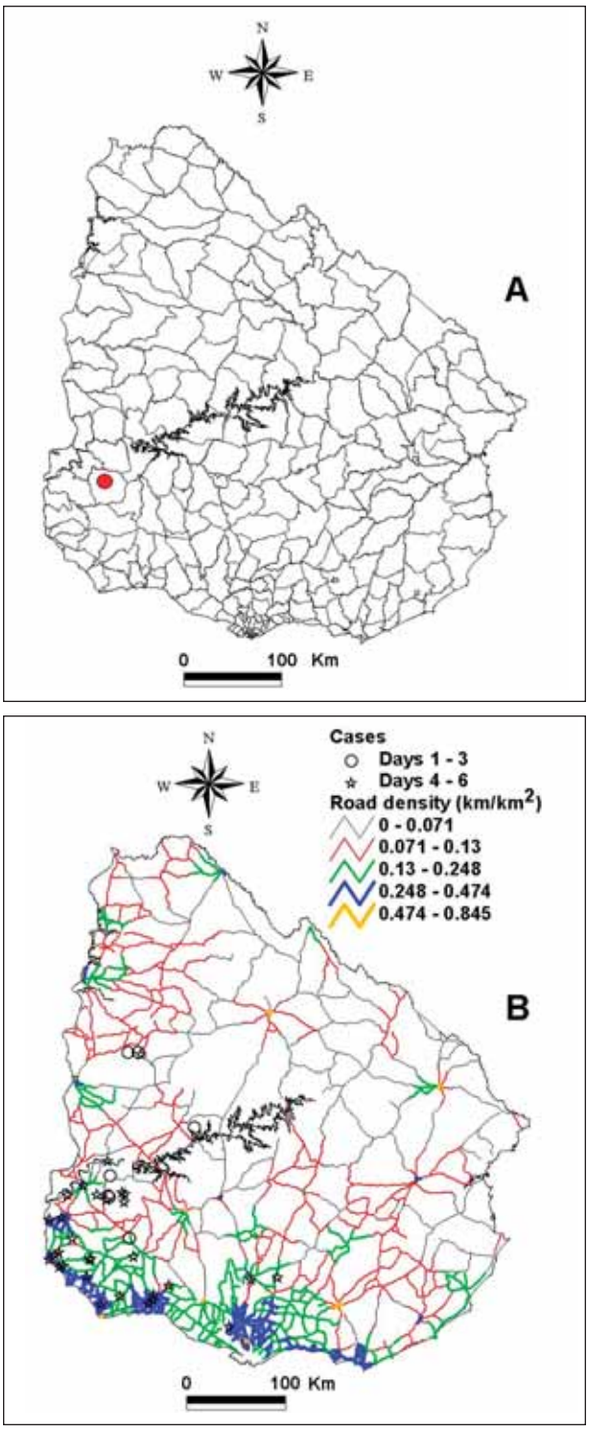

Figure 1—County location in Uruguay and site of the first herd reported as infected (red dot) during the 2001 outbreak of FMD (A) and location of farms with infected cattle during the first week of the outbreak (B).

between all centroids were calculated by use of the GIS tools, which selected a distance larger than the largest distance between any pair of points in the territory under study.

Three steps were used to generate data on road density. First, the total area of each county was determined by accessing the county value for area. Second, the national highway layer (excluding urban areas)[c] was intersected with the county layer

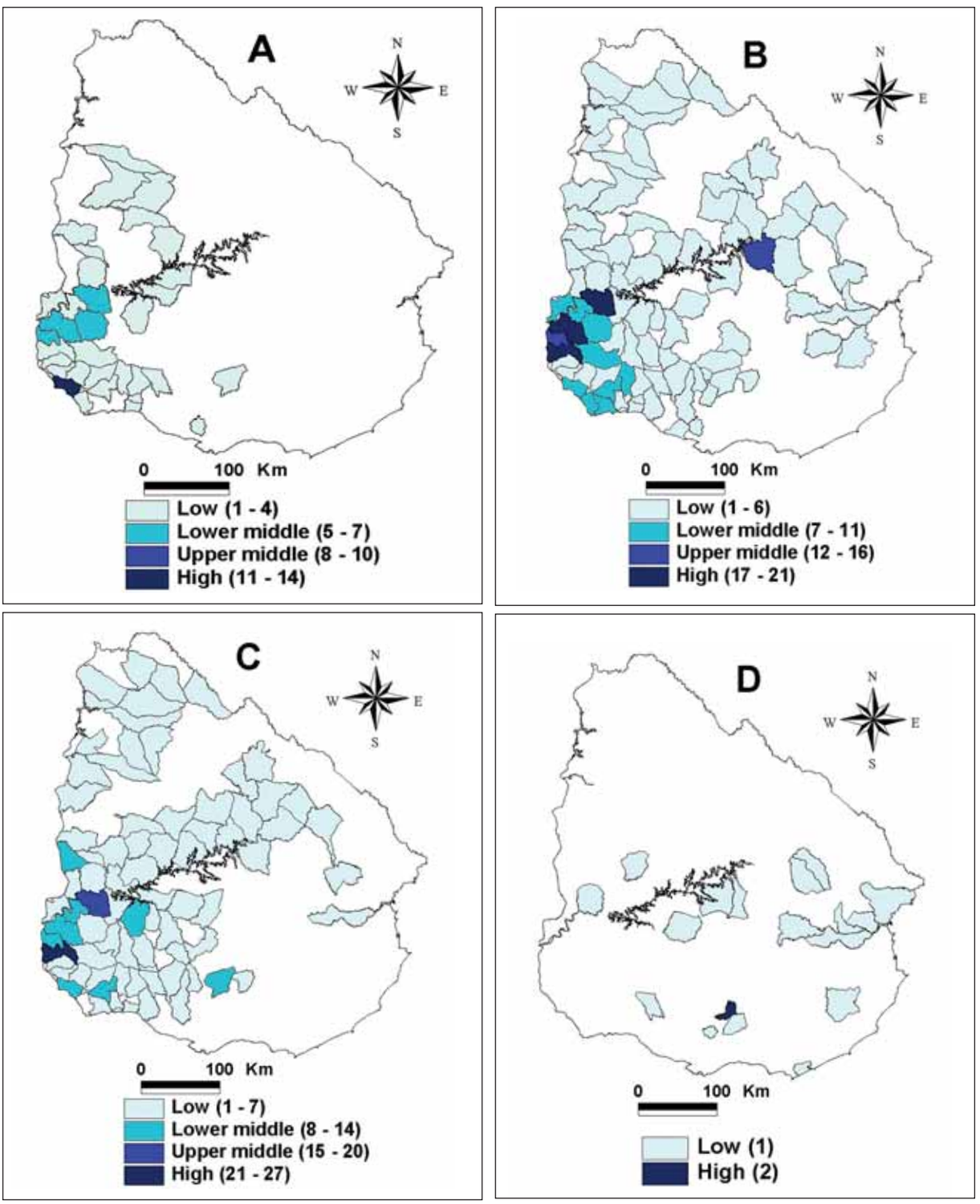

Figure 2—Number of farms with cattle infected with FMD per county at the beginning (week 1; A), peak (weeks 4 [B] and 5 [C]), and end of the 2001 epidemic (week 11; D). Number of farms with infected cattle were categorized into 4 equal-sized classes on the basis of the number of cases per week (in parentheses), except for week 11 when the number of cases was insufficient to be separated into 4 classes.

to characterize and identify road segments by county. Length of road segments was then summarized for each county (ie, the total length of roads was divided by total area of the county).

Creation of matrices and tables—The GIS-generated matrix of all pairs of intercounty (centroid-to-centroid) distances (13,203 county pairs), the table containing density of county roads, and the matrix including the number of infected cattle per week and county identifier were transferred into and processed by use of technical computing software.[d]

Spatial autocorrelation—Spatial connectivity involved Euclidean distances (ie, number of kilometers) between counties with infected cattle (distance between centroids) and road density (road distance divided by county area, a non-Euclidean distance measure). The Moran *I* coefficient was used to analyze spatial autocorrelation.[13] Positive values for spatial autocorrelation indicate that sites spatially closer to each other than the mean distance have similar numbers of cases, whereas negative values for spatial autocorrelation indicate the opposite.

The Moran *I* coefficient of autocorrelation was calculated as follows:

$$I = (n \sum_{i=1}^{n} \sum_{j=1}^{n} w_{ij}\, z_i\, z_j)/(S_0 \sum_{k=1}^{n} z_k^2)$$

where n is the number of counties, i and j are counties (i and j cannot be the same county), w_{ij} is the spatial connectivity matrix, z_i is the difference between the prevalence in county i and the overall mean prevalence, z_j is the difference between the prevalence in county j and the overall mean prevalence, S_0 is an adjustment constant, k is a county index, and z_k is the difference between the county index and overall index. In addition, $z_i = x_i - x$, where x_i is the weekly number of cases/100 farms in county i and x is the mean prevalence. The value for w_{ij} is calculated by use of the following equation:

$$w_{ij} = f\ (d_{ij}, r_i, r_j) = (d_{ij})^{-a}\ (r_i\ r_j)^b$$

where d_{ij} is the matrix of the Euclidean distance between counties i and j (i and j cannot be the same county), r_i is the road density for county i, r_j is the road density for county j, the value for variable a is a measure of the degree of epidemic diffusion in relation to distance (ie, there is greater diffusion at shorter distances),[37-41] and the value for variable b is a measure of the extent of connectivity between counties (ie, greater road density results in greater connectivity), regardless of distance. For fixed positive values of variable a, large values of variable b support local spread as well as long-distance spread because higher local road density is associated with higher interstate highway density. Values for variables a and b were estimated by maximizing the spatial autocorrelation coefficient as reported elsewhere[6] as follows:

$$I^* = \sum_{t=1}^{11} I\ (t, a, b)$$

where $a > 0$, $b > 0$, and t is time (week of the epidemic). The value for S_0 was calculated as follows:

$$S_0 = \sum_{i=1}^{n} \sum_{j=1}^{n} w_{ij}$$

where i and j cannot be the same county.

Spatial-temporal autocorrelation—Interactions of space and time were analyzed by use of the Mantel coefficient $I_{s\text{-}t}$.[14,15] The $I_{s\text{-}t}$ coefficient was calculated by use of the following equation:

$$I_{s\text{-}t} = \sum_{i=1}^{n} \sum_{j=1}^{n} w_{ij}\, y_{ij}$$

where y_{ij} indicates the closeness in time between infections and i and j cannot be the same county. The first moments of the Moran *I* and Mantel $I_{s\text{-}t}$ statistics are reported elsewhere.[6] Observations were assumed to be random independent samples from an unknown distribution function relative to the set of all possible values of *I* or $I_{s\text{-}t}$ when the x_i were randomly permuted around the county system.[6] The matrix y_{ij} was defined as $y_{ij} = 1$ when county i had values greater than the mean number of cases/100 farms (total number of susceptible farms/county) at week t and county j also had values greater than the mean number of cases/100 farms at week t – m; otherwise, y_{ij} was equal to 0. This cross-correlation at lag m measured the temporal correlation of events at time t and those at a specified preceding point (ie, m weeks earlier).

Spatial correlograms—Interaction between county pairs was measured as a function of their distance from each other as described elsewhere.[6] The graphic display of the global spatial autocorrelation coefficient (Moran *I*) plotted against the distance lag (correlogram) was determined by use of the following equation:

$$I_{(g)} = (n \sum_{i=1}^{n} \sum_{j=1}^{n} w_{ij}\, z_i\, z_j)/(S_0 \sum_{k=1}^{n} z_k^2)$$

Table 1—National weekly case prevalence during the first 11 weeks of an epidemic of FMD in Uruguay that began on April 23, 2001.

Week of the epidemic	No. of new cases*	No. of susceptible farms in counties with infected animals	Overall county herd prevalence (per 100 county farms)
1	88	4,443	1.98
2	228	11,098	2.05
3	220	10,584	2.08
4	303	12,076	2.51
5	293	10,703	2.74
6	235	12,781	1.84
7	176	11,407	1.54
8	93	8,008	1.16
9	41	4,676	0.88
10	28	3,138	0.89
11	19	2,724	0.70

*Number of farms reporting infected animals.

where g is the distance between the 2 counties, the matrix w_{ij} contains values of 1 for all the links among county pairs (i, j) located within the distance g and values of 0 for all other links not included within the Euclidean distance g, and i and j are not the same county. The temporal correlogram is the plot of $I_{s\text{-}t}$ as a function of the time lag m. Hence, the temporal correlogram was used to determine the extent of spatial-temporal autocorrelation for various time lags.

Contribution of specific links to the overall spatial autocorrelation—On the basis of network analysis, relationships between nodes (ie, counties) can be described by their links.[5,7] County pairs were considered connected by a spatial link when their contribution to the global spatial autocorrelation coefficient did not equal 0. The contribution of specific spatial links was defined as the link strength (index) between counties with infected cattle (i, j) located within a distance g, as indicated by use of the following equation:

$$I_{ij\ (g)} = (z_i\ z_j)/(\sum_{k=1}^{n} z_k^2)$$

where $I_{ij\ (g)}$ is the contribution of the specific spatial link.

Statistical analysis—Spatial-temporal autocorrelation and link indices were calculated by use of mathematical software.[d] Normality (No. of farms/county and link index, which were tested by use of the Anderson-Darling test) and comparisons among medians (assessed by use of the Mann-Whitney test) were conducted by use of a statistical program.[e] For all tests, values of $P < 0.05$ were considered significant.

Results

Epidemic scenario and model validation—The 2001 epidemic began in the southwest portion of Uruguay and reached a peak (county-level) farm prevalence at week 5 (**Table 1**). The median road density of all counties reporting infected animals during the first week was 0.24 km/km^2, which differed significantly ($P = 0.01$) from that for the remainder of the country (0.12 km/km^2; Figure 1). A dissimilar spatial pattern was observed over time (**Figure 2**; **Table 2**). The distribution of the number of susceptible farms per county did not disprove a normal distribution ($P > 0.05$; **Figure 3**). The normality assumption of the spatial autocorrelation (which requires an estimated minimum of 20 county pairs/observation) was met during at least the first 9 weeks of the epidemic because all distance lags up to approximately 440 km reported > 20 county pairs.

Spatial-temporal autocorrelation—Maximization of the spatial autocorrelation index was evident when variable a = 0.46 and variable b = 0.06 (data not shown). The Moran *I* null hypothesis (lack of spatial autocorrelation) was rejected. Until at least the sixth week of the epidemic, sites closer to each other (clusters) had significantly more infected cattle than sites located at the mean (or greater) distance from each

Table 2—Infective connectivity for county pairs containing cattle infected with FMD that had the highest index link.

Time period and distance	County pairs	Infective link index*	County connecting with ≥ 2 other counties through a high index link	No. of links†
Before vaccination and <100 km between county pairs‡	409, 1704	3.07	409	7
	409, 1709	2.48	1704	4
	409, 1707	2.02	1707	2
	407, 409	1.91	1709	2
	1704, 1709	1.81	407	2
	409, 1705	1.66	NA	NA
	409, 412	1.40	NA	NA
	409, 1708	1.33	NA	NA
	407, 1704	1.32	NA	NA
	1704, 1707	1.30	NA	NA
After vaccination and < 100 km between county pairs§	1707, 1709	2.54	1709	6
	1705, 1709	2.14	1704	3
	1704, 1709	2.05	1707	3
	1704, 1707	1.58	1705‖	3
	1705, 1707	1.49	NA	NA
	1703, 1709	1.36	NA	NA
	414, 1709	1.34	NA	NA
	409, 1709	1.17	NA	NA
	1704, 1705	1.15	NA	NA
Before vaccination and > 400 km between county pairs¶	105, 409	3.37	409#	1
	105, 407	2.17	407#	1

*Percentage of the overall spatial autocorrelation index explained by a specified spatial infective link index connecting 2 counties (1 is assumed to be the infector and the other is assumed to be the target). †Counties with ≥ 2 links (both of which had high indices) are regarded to possess greater potential for epidemic spread (infector site), whereas those observed with only 1 link or observed at a later time during the epidemic are regarded as target sites. ‡Represents weeks 1 and 2 during the epidemic for 2,308 spatial links with a mean ± SD link index of 0.043 ± 0.15. §Represents weeks 3 through 11 during the epidemic for 2,151 spatial links with a mean ± SD link index of 0.046 ± 0.14. ‖County No. 1705 did not appear to have links by itself because all 3 links to it are explained by links for counties Nos. 1704, 1707, and 1709. ¶Represents weeks 1 and 2 during the epidemic for 394 spatial links with a mean ± SD link index of 0.254 ± 0.23. #Because counties Nos. 407 and 409 already contained infected cattle at week 1 and county No. 105 did not report infected cattle until week 5, these connections appear to rule out county No. 105 as the site that infected counties Nos. 407 and 409.

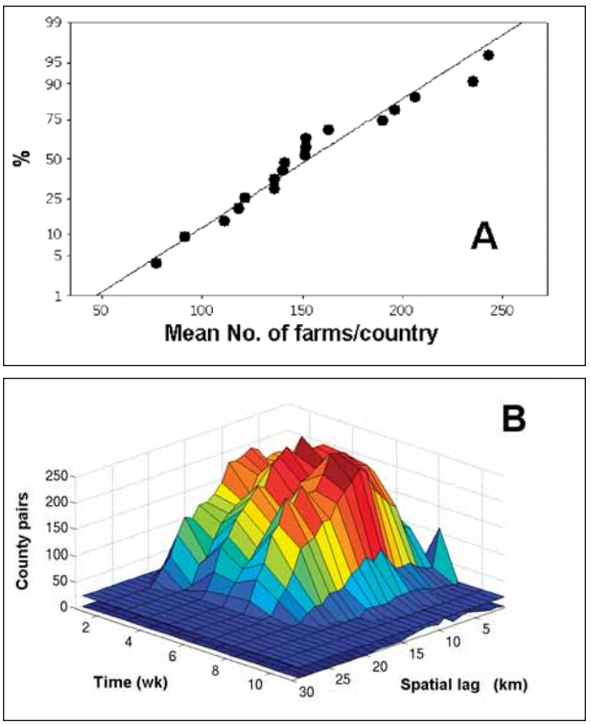

Figure 3—Distribution of the national number of total (susceptible) farms per county (aggregated at the state level; n = 18 states; A) and the number of observations for county pairs that contained infected cattle at specific time points (weeks during the outbreak) or distance lags (between county pairs; B). Notice that the number of susceptible farms has approximately a normal distribution ($P = 0.410$; Anderson test) with a mean ± SD of 153.3 ± 45.7 farms/county. The number of county pairs exceeded 20 for each of the first 9 weeks of the epidemic.

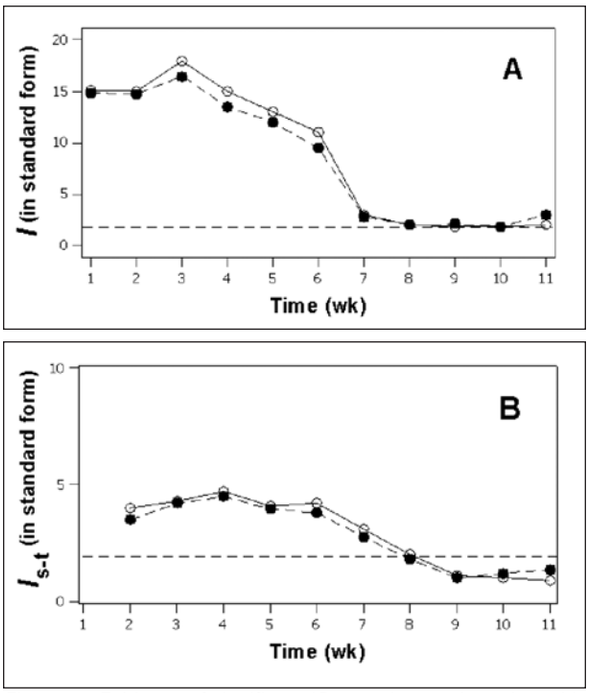

Figure 4—Evidence of significant ($P < 0.05$) case clustering with spatial autocorrelation (Moran *I*; A) and spatial-temporal autocorrelation (Mantel I_{s-t}; B) observed during the first 6 weeks of the 11-week epidemic of FMD. Tests were conducted considering case prevalence/100 susceptible county farms (open circles, solid lines) and without considering susceptible county farms (closed circles, broken lines). The upper limit (2 SDs) for values not autocorrelated is indicated in each graph (dashed line).

other (**Figure 4**). In addition, analysis of the Mantel I_{s-t} indicated that in weeks 1 through 6, spatial clusters were associated with time because adjacent sites had significantly more infected cattle at shorter time periods than sites more distant in time and place. Because exotic diseases have zero prevalence before an outbreak and every infection needs to be controlled (regardless of the size of the susceptible population), Mantel and Moran tests were also calculated without considering the total size of the susceptible population, and both calculations yielded similar results.

Correlogram analysis—Analysis of spatial correlograms (conducted before and after vaccination was implemented) indicated a significant positive autocorrelation among county pairs with infected animals located within approximately 120 km from each other for weeks 1 and 2 of the outbreak and within 80 km of each other for weeks 3 through 11. A significant negative spatial autocorrelation was observed for county pairs with infected cattle located 120 to 400 km from each other only at weeks 1 and 2 of the outbreak. A second cluster, which was not significant, was evident for county pairs with infected cattle located > 400 km from each other (**Figure 5**). The temporal correlogram indicated significant temporal-spatial autocorrelation for time lags of up to 3 weeks ($m < 4$). When specific weeks were considered, spatial correlograms did not reveal regional effects. During the first 6 weeks of the epidemic, significant positive spatial autocorrelation was observed each week for county pairs with infected cattle located within 120 km of each other, whereas a significant negative autocorrelation lasted for at least the first 5 weeks for cattle located at 300-400km (**Figure 6**).

Infective spatial connectivity—Analysis of infective link indices (percentage of the overall spatial autocorrelation explained by specific infective links) revealed a clear departure from normality (**Figures 7–9**). County pairs with infected cattle located < 120 km from each other during weeks 1 and 2 had 10 links (including 5 different counties) with indices substantially higher than the mean. Three of those 5 counties also had the highest link indices at weeks 3 through 11. The remaining 2 counties were involved in significant long-distance links for weeks 1 and 2, and analysis also suggested that they departed from normality, but not significantly, for weeks 3 through 11 (Table 2).

Categorization of infected sites by link index, time, and number of links—Analysis of the data suggested 3 classes of counties in terms of potential disease dispersal during the epidemic. The first class included 5 counties in which infected cattle were observed within the first 3 days of the epidemic (minimal time compatible with a replication cycle of the infective agent; hence, possible primary cases; Figure 7). All of these counties, except for 1, had low index links. The second class included 5 counties that

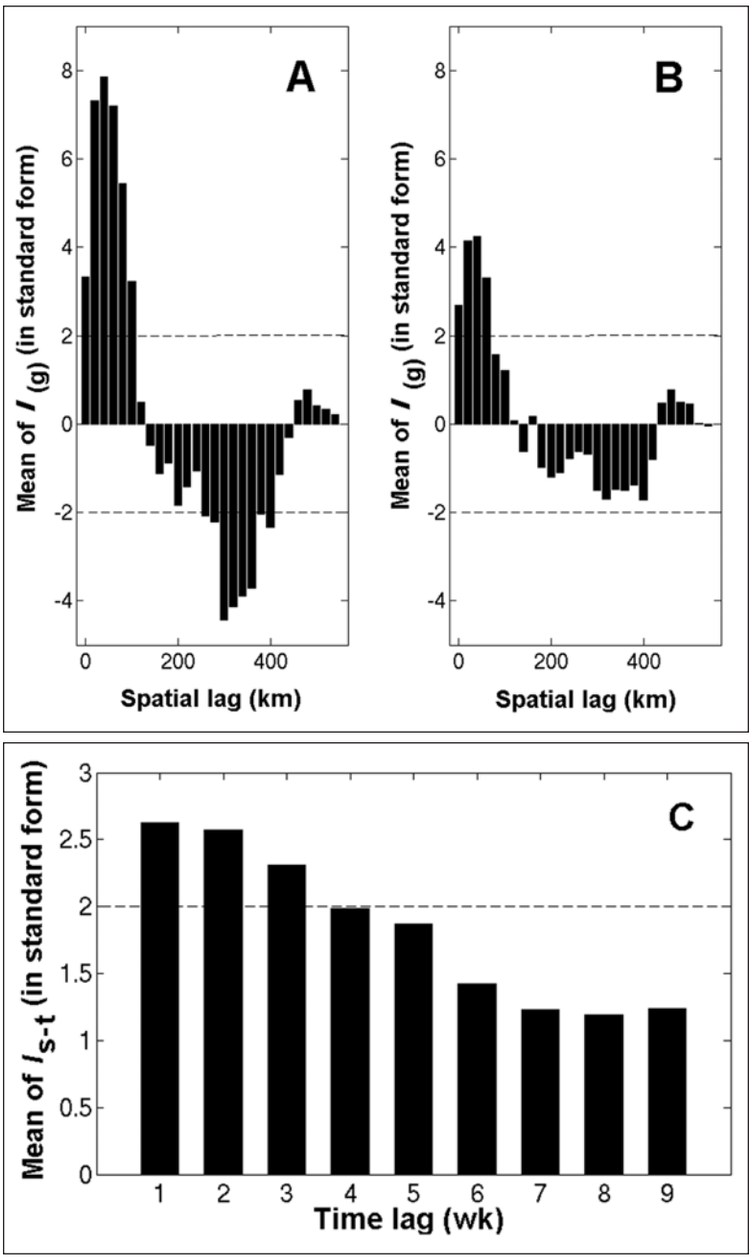

Figure 5—Mean spatial correlograms for the periods during the epidemic before vaccination (weeks 1 and 2; A) and after vaccination (weeks 3 through 11; B) and the temporal correlogram for the entire 11 weeks of the epidemic (C). Correlograms express the mean autocorrelation index ($I_{(g)}$) for infected sites located between the specified distance lag (each bar corresponds to a distance lag of 20 km). For example, the mean index for county pairs that contained infected cattle in which the counties were located at a distance of 40 to 60 km from each other was 7.9 (the third bar) for the period before vaccination. Correlograms reveal a significant ($P < 0.05$) positive autocorrelation for county pairs that contained infected cattle in which the counties were located up to approximately 120 km apart, followed by a significant negative autocorrelation (120 to 400 km between county pairs at weeks 1 and 2) and a second but not significant cluster at > 400 km between county pairs before and after interventions. The temporal correlogram reveals that case prevalence was significantly correlated for a time lag of up to 3 weeks. The upper and lower limits (± 2 SDs) for values not autocorrelated are indicated in each correlogram (dashed lines).

had the highest index links connecting with ≥ 2 other counties. One of the counties was possibly a primary site (with infected animals reported within 3 days of the outbreak), whereas the other 4 counties all reported infected cattle within 4 to 6 days of the epidemic. These counties had both short- and long-distance connections. The third class involved counties reporting infections after week 1 of the epidemic and had mean link indices (counties regarded as targets). When 2 counties were connected, time during the epidemic helped to generate hypotheses that distinguished the putative infector (earlier case) from the putative infected (later case [target]; Figure 9; Table 2). When 1 county of the pair connected by a high index link was involved in multiple links, but the other county was not, the first county was hypothesized to be the infector (**Table 3**).

Applications in the decision-making process—All counties but one of those reporting primary cases did not appear to facilitate spread of the disease during the epidemic. Four of 5 counties that had the highest link indices and connected with at least 2 other counties had 2.5 times as many cases by week 11 as 4 of 5 counties that contained cattle infected during days 1 to 3 of the epidemic. The second group of counties (counties with a high index link) reported their first infected animal on days 4 to 6 of the epidemic (time frame compatible with a secondary infection); when combined with another high index link county that reported an infected animal at day 1 to 3, this provided a county median of 0.073 cases/km^2 by week 11, whereas the remaining counties reporting cases at days 1 to 3 (none of which were high index link counties) had significantly ($P = 0.02$; Mann-Whitney test) fewer infected cattle (county median, 0.027 cases/km^2) by week 11 (Table 3). Counties with a high index link (n = 5) also had a significantly ($P = 0.01$) higher median road density (0.26 km/km^2), compared with the 271 other counties with infected cattle (0.126 km/km^2).

Discussion

Because observational epidemiologic analyses do not allow experimental designs, theories can only use historical data to attempt validation. However, such data may possess unknown sources of bias or lack critical variables. For example, the number of farms considered in the study reported here was based on the 2000 Agricultural Census, a data set not necessarily applicable for the study of this epidemic. Accordingly, the model described should not be perceived as an analysis of the FMD epidemic that took place in Uruguay in 2001 but, instead, as an eval-

Table 3—Comparison of control efficacy for an outbreak of FMD on the basis of spatial-based versus traditional approaches.

	Spatial-based approach*				Traditional approach†			
County No.	**County Spatial links**	**County area (km^2)**	**All cases reported through week 11**	**Cases/km^2 through week 11**	**Primary county No.**	**Primary county area (km^2)**	**All cases reported in primary counties through week 11**	**Cases/km^2 in primary counties through week 11**
407	2	382.0	28	0.073	1108	2,262.2	13	0.006
409	7	474.0	72	0.152	1209	1,294.2	8	0.006
1704	4	1,070.2	70	0.065	1706	1,176.9	37	0.031
1707	2	1,047.8	69	0.066	1707‡	1,047.8	69	0.027
1709	2	763.8	93	0.122	1708	1,218.9	33	0.066
Total§	NA	**3,737.8**	**332**	**0.478**	NA	**6,999.9**	**160**	**0.136**
Median‖	NA	905.8	NA	0.073	NA	1,256.6	NA	0.027

*Counties with a high index link (sufficient counties) are those that have substantially high infective connectivity indices (at least 3.5 times greater than 2 SDs), link with at least 2 other counties, and report infected cattle earlier than the other county sharing the infective link. †Counties without a high index link (necessary counties) are those that report infected cattle during the first 3 days of the epidemic (minimal time for the replication cycle of FMD virus) and hence are hypothesized to be primary cases and also have link indices within the mean + 2 SDs. ‡County No. 1707 is a county with a high index link that reported infected cattle during the first 3 days of the epidemic (primary cases). §Expressed in percentages, counties with a high index link reported > 2 times as many cases (332/160 [207.5%]) as counties without a high index link. Expressed as area, total surface for counties with a high index link represented almost half that for counties without a high index link (3,737.8 km^2/7,000.0 km^2 [53.4%]). Expressed as total number of cases prevented per km^2, a control campaign implemented in counties with a high index link could have prevented 3.5 times more cases per square kilometer than a similar campaign implemented in counties without a high index link (0.478/0.136 = 3.51). ‖Expressed as median number of cases prevented per county, a control campaign implemented in counties with a high index link could have prevented 0.073 cases/km^2, which was significantly ($P = 0.02$; Mann-Whitney test) higher than the number of cases prevented per county (0.027 cases/km^2) had the same control campaign been implemented in counties without a high index link.

NA = Not applicable.

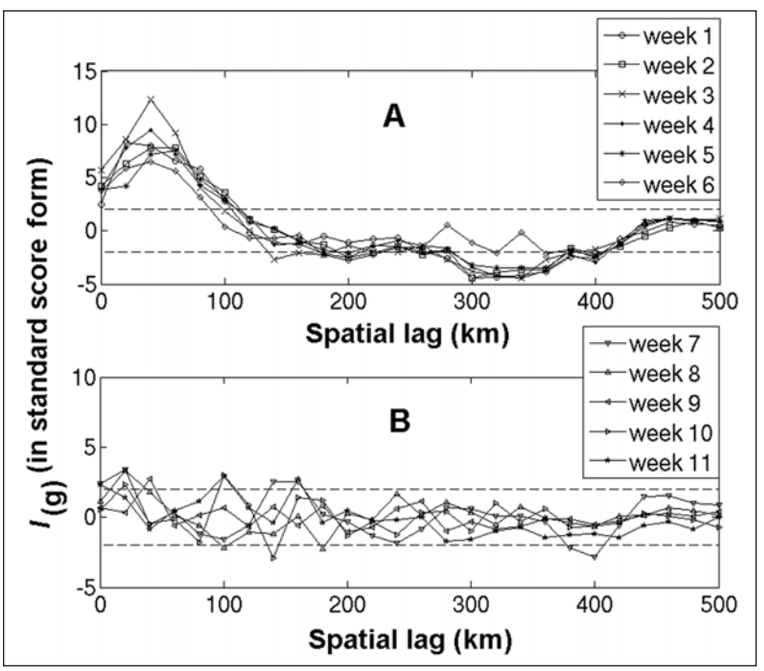

Figure 6—Spatial correlograms calculated for weeks 1 through 6 (A) and 7 through 11 (B) of the epidemic. There is a lack of evidence of regional effects on case clustering during the first 9 weeks of the epidemic. Each correlogram reveals a significant ($P < 0.05$) positive autocorrelation (values > 2 SDs [upper dashed line]) within approximately 120 km between county pairs and a negative autocorrelation (values < 2 SDs [lower dashed line]) between county pairs approximately 280 and 400 km apart.

uation of a spatial method that uses a hypothetical (although realistic) scenario for the epidemic. Despite that caveat, the analysis of assumptions on which spatial autocorrelation was based revealed adequate sample size (> 20 county pairs/observation) and no departure from normality.[29] Two measures of spatial-temporal autocorrelation (with and without consideration of denominator data) yielded similar results. Similar week-specific correlograms suggested that delayed reporting did not bias these findings. The use of Euclidean and non-Euclidean distances was justified by the fact that there was a maximized spatial autocorrelation index when variable a = 0.46 and variable b = 0.06.[6]

Significant positive (< 120 km between counties with infected animals) and negative (> 120 but < 400 km between counties with infected animals) spatial autocorrelations were observed every week for at least the first 5 weeks (Figure 6). Such findings suggested that, once structured, the epidemic network was rather robust and static. Three major spatial autocorrelation patterns have been described[42]: a monotonic decreasing pattern (a positive-only significant autocorrelation without a significant negative autocorrelation; also known as a patchy pattern); a bimodal pattern characterized by significant positive spatial autocorrelation for short-distance lags, followed by significant negative spatial autocorrelation for long-distance lags, as was evident in the study reported here; and lack of spatial pat-

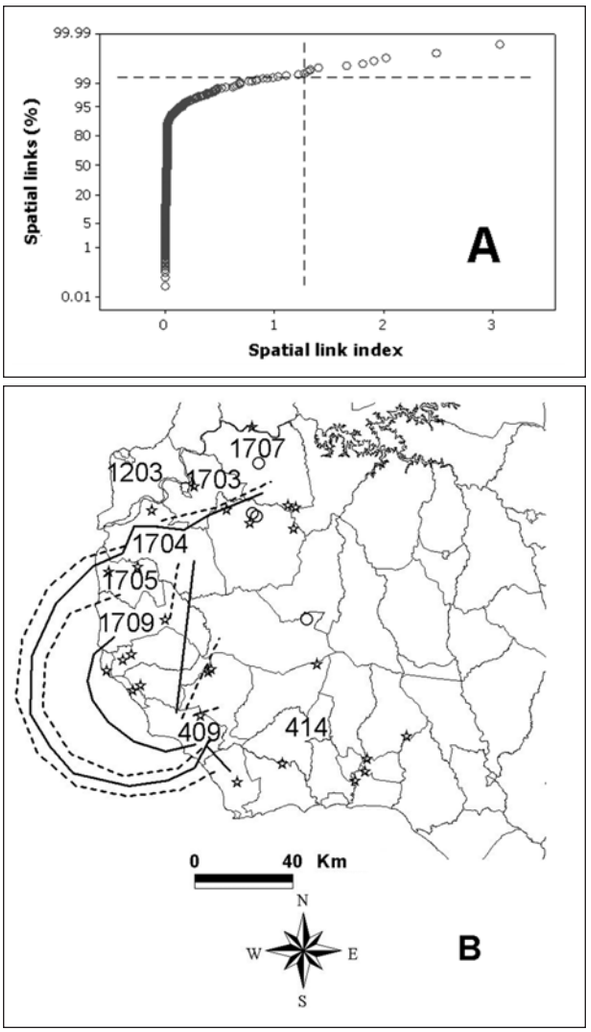

Figure 7—Contributions of specific links between county pairs that contained infected cattle to the overall autocorrelation index for the period before vaccination (weeks 1 and 2) for county pairs located < 120 km apart (A) and a map of the southwestern region of Uruguay indicating the 10 highest spatial infective link indices (lines) between county pairs (B). In panel A, 10 links involving 5 different counties have a non-normal distribution ($P > 0.05$; Anderson test) and substantially higher indices than the mean + 3 SD (upper right quadrant for the dashed lines). The link index, number of high index links per county, and time at which cattle in the county became infected help to distinguish infector from infected counties (Table 3). Solid and broken lines represent high index links and low index links, respectively (B). *See* Figure 1 for remainder of key.

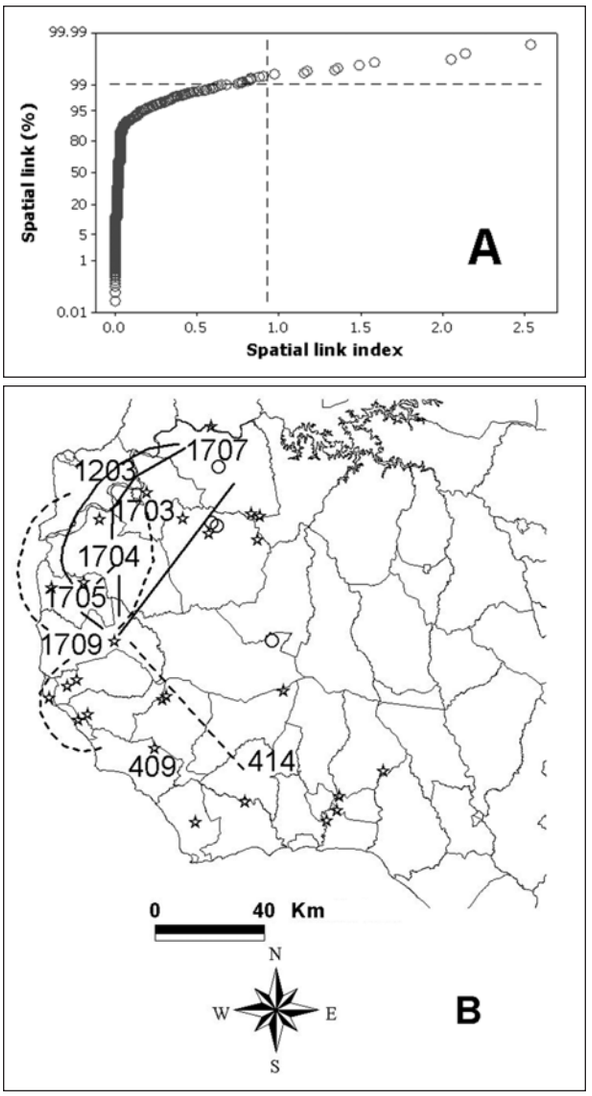

Figure 8—Contributions of specific links between county pairs that contained infected cattle to the overall autocorrelation index for the period after vaccination (weeks 3 through 11) for county pairs located < 120 km apart (A) and a map of the southwestern region of Uruguay indicating the 10 highest spatial infective link indices (lines) between county pairs (B). *See* Figures 1 and 7 for remainder of key.

terns (when the Moran *I* coefficient is not significant). Although monotonic and decreasing Moran indices (eg, lacking a significant negative autocorrelation) are usually found in other fields, negative structures are not rare in epidemiologic investigations.[29] Possible causes of significant negative autocorrelations include poor local connectivity for 1 member of county pairs (eg, lower road density, factor associated with lower farm density, or fewer adjacent farms).[24,25] A correlogram pattern with significant positive and negative autocorrelations for short- and long-distance lags, respectively, can be interpreted as a linear gradient at macroscales such that when 1 member of the pair is situated farther than a certain critical distance from the other member of the pair, case prevalence typically has opposite values.[42] Nonsignificant links at even greater distances for lags (> 400 km) resembled small-world–like connections.[5] As indicated by the lack of significance, such connections do not necessarily result in additional disease spread during an epidemic because local conditions (ie, poorer local connectivity) may prevent viral dispersal.

Spatial analysis facilitated data-driven generation of hypotheses. Counties with infected cattle could be categorized as possessing greater potential for disease dispersal during the epidemic on the basis of 3 criteria (having a high index link [ie, to be an outlier or county with a high index link], connecting with ≥ 2 other counties, and reporting infections before the other member of the pair). Counties reporting infections on days 1 to 3 of the outbreak (primary cases) were

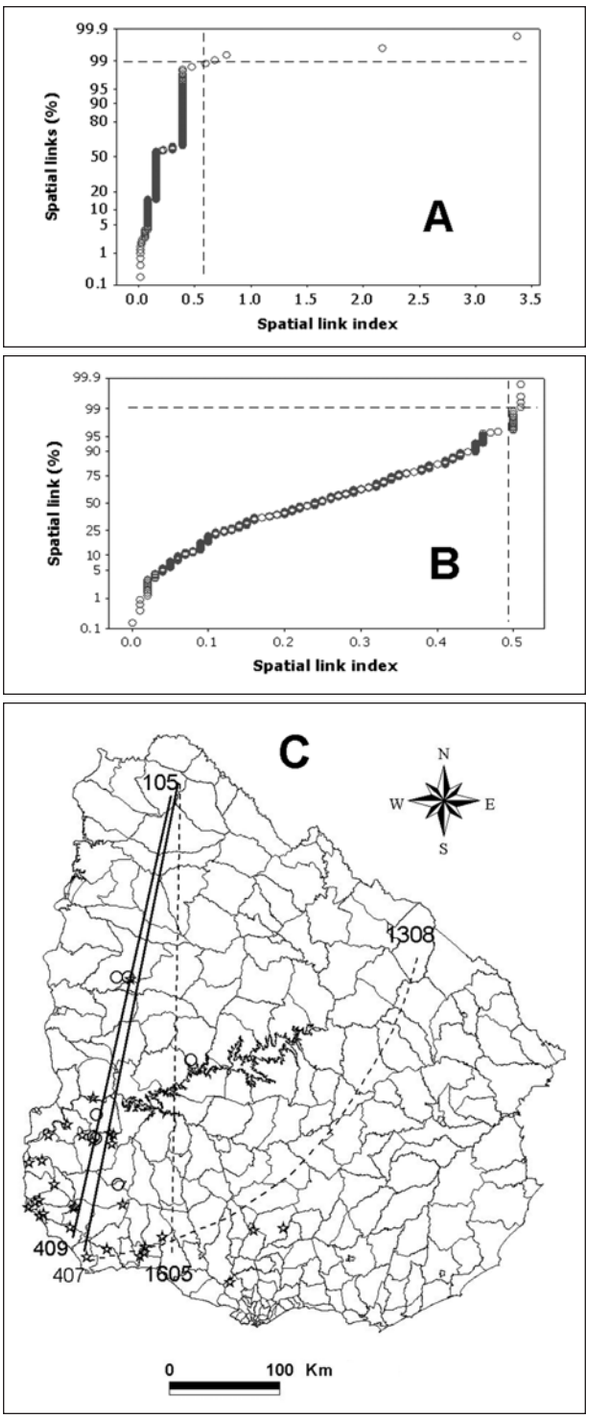

Figure 9—Contributions of specific links between county pairs that contained infected cattle to the overall autocorrelation index for the period before vaccination (weeks 1 and 2; A) and after vaccination (weeks 3 through 11; B) for county pairs located > 400 km apart and a map of Uruguay that indicates the 4 highest intercounty link indices (lines) before vaccination (C). A significant ($P < 0.05$) departure from normality is observed before but not after vaccination was implemented. The highest link indices involve 2 (putative) target counties (Nos. 105 and 1308). Despite these substantially high link indices (which were approx 2 SDs), the global spatial test did not indicate significant autocorrelation before vaccination for county pairs containing infected cattle that were located > 400 km apart. *See* Figures 1 and 7 for remainder of key.

regarded as necessary sites, whereas those displaying higher index links (and connecting with at least 2 additional counties) were hypothesized to possess greater risk for other counties (sufficient cause of disease spread during the epidemic). Counties paired with those that had sufficient cause of disease spread were suspected to be target sites. This working hypothesis distinguished counties infected first (necessary causes, although not necessarily the cause of disease spread) from those that had a high index link (ie, those hypothesized to seed new cases into target sites), regardless of when and where they got the infection. This conceptualization is similar to that of a model in which it was proposed that spatial features result in differing diffusion models during an epidemic.[40] Although daily data on time of detection of infected animals facilitate the richest generation of hypotheses, even when such data are not available or are available but not used because of possible errors (eg, delayed reporting and underreporting), information on link indices alone identifies county pairs that have indices much higher than the mean (outliers suspected to influence disease dispersal).

Although other factors associated with disease spread during an epidemic (ie, markets) cannot be ruled out, spatial analysis may generate evidence of case clustering, whether there are short- or long-distance connections (or both), and whether there are changes in location of cases over time in relation to interventions. Identification of infected sites with greater epidemic risk (counties with a high index link) did not support the hypothesis that all infected cattle had equal influence on disease spread nor the theory of homogeneous mixing, which assumes that all susceptible and infected cattle are located at similar distances from each other and possess similar risk for becoming infected or for infecting others.[40] This theory results in undifferentiated control policies, such as implementation of buffer rings (ie, regional circles of fixed diameter within which the same control policy is conducted).[43] The fact that the first county with infected cattle and 3 other counties in which there were primary infections apparently failed to promote disease spread also argued against the homogeneous mixing theory.

Spatially explicit assessment of infective connectivity may be applied to evaluate control policy. For example, when only 2 time periods were considered, spatial autocorrelation analysis revealed a reduction of approximately 40 km in the mean distance between counties for the cluster (from 120 km at weeks 1 and 2 to 80 km at weeks 3 through 11), which supports the hypothesis that vaccination reduced disease spread during the epidemic. However, evaluation of week-specific correlograms did not reveal evidence of regional differences up to week 6 of the epidemic, which suggests that the 40-km reduction may reflect the end of the epidemic (when many counties did not report cases). These results may support the hypothesis that the conclusion of the epidemic was attributable to several factors, including lack of susceptible herds and a ban on animal movement that was imposed in week 1.

The approach described here was also informative, facilitating the explanation of apparent contradictions.

Although a second cluster was suggested by correlograms for sites located at > 400 km between counties with infected cattle before and after vaccination was conducted, which is in agreement with the expected limited disease dispersal for infected animals located at the edge of the territory being infected,[40] the cluster at > 400 km was not significant (Figures 5, 6, and 9). However, at weeks 1 and 2, link analysis identified 2 counties that had a high index and long-distance connections. The contradiction between (global) correlogram analysis and link analysis may be explained once local factors are considered (ie, edge effects and a lower density of local roads in target counties connected by long-distance links may prevent further disease dispersal because there is poor local connectivity).

Cost-benefit analysis may also be generated by the approach used in the study reported here. Had a policy focusing on all counties reporting primary cases been adopted (on the basis of the theory that all cases equally contribute to disease spread during an epidemic), it may have been inefficient and insufficient. In contrast, a policy focused on high-index link counties could have been 2.5 to 3 times more beneficial than undifferentiated control policies (Table 3). Observations of significant case clustering and significant negative autocorrelation (for counties located > 120 to < 400 km between counties with infected cattle), noticed as early as week 2 (when vaccination had not been implemented), could have led to differentiated control measures (ie, regionalization).[44]

Infective link analysis can be interpreted by considering epidemics as processes that connect at least 2 points through a line. The local Moran test has been used[12,45,46] to focus on the contribution of each point to the overall (global) spatial autocorrelation. In contrast, the method described here focused on the line connecting the 2 points. Although local Moran tests assess inputs and outputs, infective connectivity emphasizes the intermediate process that takes place at some time point before the outcome is noticed. Such emphasis informs on earlier phenomena, which can be used to generate hypotheses on factors facilitating (or preventing) disease dispersal during an epidemic and possibly to identify case clustering in adjacent sites and in sites located far apart from each other. When based on data of a smaller scale (ie, farm-level data), spatial autocorrelation and link analysis may facilitate real-time control of rapidly disseminated diseases.

a. Arc View GIS 3.3, ESRI, Redlands, Calif.
b. Arc View 8.0, ESRI, Redlands, Calif.
c. Geographic Service, Ministry of Defense, Montevideo, Uruguay.
d. Matlab, Mathworks Inc, Natick, Mass.
e. Minitab 14, Minitab, State College, Pa.

References

1. Rainham DGC. Ecological complexity and West Nile Virus—perspectives on improving public health response. *Can J Public Health* 2005;96:37–40.

2. Langlois JP, Fahrig L, Merriam G, et al. Landscape structure influences continental distribution of hantavirus in deer mice. *Landscape Ecol* 2001;16:255–266.

3. Wilesmith JW, Stevenson MA, King CB, et al. Spatio-temporal epidemiology of foot-and-mouth disease in two counties of Great Britain in 2001. *Prev Vet Med* 2003;61:157–170.

4. Milgram S. Small-world problem. *Psychol Today* 1967;1:61–67.

5. Watts DJ, Strogatz SH. Collective dynamics of 'small-world' networks. *Nature* 1998;393:440–442.

6. Cliff AD, Ord JK. Measures of autocorrelation in the plane; and Distribution theory for the join-count, I, and c statistics. In: Cliff AD, Ord JK, eds. *Spatial processes: models and applications*. London: Pion Ltd, 1981;1–65.

7. Bollobás B. Models of random graphs. In: Bollobás B, Fulton W, Katok A, et al, eds. *Random graphs*. Cambridge Studies in Avanced Mathematics 73. Cambridge, UK: Cambridge University Press, 2001;34–50.

8. Morris RS, Wilesmith JW, Stern MW, et al. Predictive spatial modelling of alternative control strategies for the foot-and-mouth disease epidemic in Great Britain, 2001. *Vet Rec* 2001;149:137–144.

9. Jules ES, Kauffman MJ, Ritts WD, et al. Spread of an invasive pathogen over a variable landscape: a nonnative root rot on Port Orford cedar. *Ecology* 2002;83:3167–3181.

10. Hawbaker TJ, Radeloff VC. Roads and landscape pattern in northern Wisconsin based on a comparison of four road data sources. *Conserv Biol* 2004;18:1233–1244.

11. Lam NSN, Fan M, Liu KB. Spatial-temporal spread of the AIDS epidemic, 1982–1990: a correlogram analysis of four regions of the United States. *Geogr Anal* 1996;28:93–107.

12. Cocu N, Harrington R, Hulle M, et al. Spatial autocorrelation as a tool for identifying the geographical patterns of aphid annual abundance. *Agric Forest Entomol* 2005;7:31–43.

13. Moran PAP. Notes on continuous stochastic phenomena. *Biometrika* 1950;37:17–23.

14. Knox EG. The detection of space-time interactions. *J Appl Stat* 1964;13:25–29.

15. Mantel N. The detection of disease clustering and a generalized regression approach. *Cancer Res* 1967;27:209–220.

16. Jacquez GM. A κ-nearest neighbour test for space-time interaction. *Stat Med* 1996;15:1935–1949.

17. Baker RD. Testing for space-time clusters of unknown size. *J Appl Stat* 1996;23:543–554.

18. Norström M, Pfeiffer DU, Jarp J. A space-time cluster investigation of an outbreak of acute respiratory disease in Norwegian cattle herds. *Prev Vet Med* 2000;47:107–119.

19. Turnbull B, Iwano EJ, Burnett WS, et al. Monitoring for clusters of disease: application in leukemia incidence in upstate New York. *Am J Epidemiol* 1990;132(suppl 1):S136–S143.

20. Kulldorff M, Athas WF, Feuer EJ, et al. Evaluating cluster alarms: a space-time scan statistic and brain cancer in Los Alamos, New Mexico. *Am J Public Health* 1998;88:1377–1380.

21. Patil GP, Taillie C. Upper level set scan statistic for detecting arbitrarily shaped hotspots. *Environ Ecol Stat* 2004;11:183–197.

22. Tango T, Takahashi K. A flexibly shaped spatial scan statistic for detecting clusters. *Int J Health Geogr* 2005;4:11. Available at: www.ijhealthgeographics.com/content/4/1/11. Accessed June 15, 2005.

23. Rivas AL, Tennenbaum SE, Aparicio JP, et al. Critical response time (time available to implement effective measures for epidemic control): model building and evaluation. *Can J Vet Res* 2003;67:307–315.

24. Rivas AL, Smith SD, Sullivan PJ, et al. Identification of geographic factors associated with early spread of foot-and-mouth disease. *Am J Vet Res* 2003;64:1519–1527.

25. Rivas AL, Schwager SJ, Smith S, et al. Early and cost-effective identification of high risk/priority control areas in foot-and-mouth disease epidemics. *J Vet Med B Infect Dis Vet Public Health* 2004;51:263–271.

26. Alexandersen S, Quan M, Murphy C, et al. Studies of quantitative parameters of virus excretion and transmission in pigs and cattle experimentally infected with foot-and-mouth disease virus. *J Comp Pathol* 2003;129:268–282.

27. Keeling MJ, Woolhouse MEJ, Shaw DJ, et al. Dynamics of the 2001 UK foot and mouth epidemic: stochastic dispersal in a heterogeneous landscape. *Science* 2001;294:813–817.

28. Durr PA, Froggatt AEA. How best to geo-reference farms? A case study from Cornwall, England. *Prev Vet Med* 2002;56:51–62.

29. Glavanakov S, White DJ, Caraco T, et al. Lyme disease in New York state: spatial pattern at a regional scale. *Am J Trop Med Hyg* 2001;65:538–545.

30. Kao RR. The role of mathematical modelling in the control of the 2001 FMD epidemic in the UK. *Trends Microbiol* 2002;10:279–286.

31. European Commission—Health and Consumer Protection Directorate-General. *Final report of a mission carried out in Uruguay from 25 to 29 June 2001 in order to evaluate the situation with regard to outbreaks of foot and mouth disease. DG(SANCO)/3342/2001*. Brussels: European Commission, 2001. Available at: europa.eu.int/comm/food/fs/inspections/vi/reports/uruguay/vi_rep_urug_3342-2001_en.pdf. Accessed Aug 26, 2005.

32. European Commission—Health and Consumer Protection Directorate-General. *Final report of a mission carried out in Uruguay from 1 to 4 October 2001 in order to evaluate the controls in place over foot and mouth disease. DG(SANCO)/3456/2001*. Brussels: European Commission, 2001. Available at: europa.eu.int/comm/food/fs/inspections/vi/reports/uruguay/vi_rep_urug_3456-2001_en.pdf. Accessed Aug 26, 2005.

33. Doel TR. FMD vaccines. *Virus Res* 2003;91:81–99.

34. Ministry of Agriculture, Livestock and Fisheries (MGAP). MGAP home page. Montevideo, Uruguay. Available at: www.mgap.gub.uy. Accessed Jul 15, 2001.

35. Ministry of Agriculture, Livestock and Fisheries (MGAP). Directory of Agricultural Statistics. 2000–2003 annals [database online]. Montevideo, Uruguay. Available at: ww.mgap.gub.uy/diea/Anuario2003/Default.htm. Accessed Sep 10, 2005.

36. Ministry of Agriculture, Livestock and Fisheries (MGAP). Directory of Agricultural Statistics. 2003 annals [database online]. Montevideo, Uruguay. Available at: www.mgap.gub.uy/diea/Anuario2003/. Accessed Sep 9, 2005.

37. Ministry of Agriculture, Livestock and Fisheries (MGAP). Directory of Agricultural Statistics. 2000 agricultural census [database online]. Montevideo, Uruguay. Available at: www.mgap.gub.uy/Diea/CENSO2000/censo_general_agropecuario_2000.htm. Accessed Aug 26, 2005.

38. Murray GD, Cliff AD. A stochastic model for measles epidemics in a multi-region setting. *Trans Inst Br Geogr* 1975;2:158–174.

39. Hanski I. Metapopulation dynamics. *Nature* 1998;396:41–49.

40. Filipe JAN, Maule MM. Effects of dispersal mechanisms on spatio-temporal development of epidemics. *J Theor Biol* 2004;226:125–141.

41. Xia Y, Bjørnstad ON, Grenfell BT. Measles metapopulation dynamics: a gravity model for epidemiological coupling and dynamics. *Am Nat* 2004;164:267–281.

42. Felizola Diniz-Filho JA, Bini LM, Hawkins BA. Spatial autocorrelation and red herrings in geographical ecology. *Global Ecol Biogeogr* 2003;12:53–64.

43. Müller J, Schönfisch B, Kirkilionis M. Ring vaccination. *J Math Biol* 2000;41:143–171.

44. Tinline RR, MacInnes CD. Ecogeographic patterns of rabies in southern Ontario based on time series analysis. *J Wildl Dis* 2004;40:212–221.

45. Getis A, Ord JK. The analysis of spatial association by use of distance statistics. *Geogr Anal* 1992;24:189–206.

46. Anselin L. Local indicators of spatial association—LISA. *Geogr Anal* 1995;27:93–115.