Factors associated with spatial clustering of foot-and-mouth disease in Nepal

Bimal K. Chhetri · Andres M. Perez · Mark C. Thurmond

Accepted: 18 March 2010 / Published online: 7 July 2010 © The Author(s) 2010. This article is published with open access at Springerlink.com

Abstract The purpose of this study was to quantify associations between hypothesized epidemiological factors and the spatial distribution of foot-and-mouth disease (FMD) in Nepal. Spatial clustering of reports of at least one FMD case by Village Development Committee (VDC) in 2004 was examined by use of the spatial scan statistic. A Bayesian Poisson multivariate regression model was used to quantify the association between the number of reports and 25 factors hypothesized to be associated with FMD risk. The spatial scan statistic identified $P<0.01$ two clusters of FMD reports. Large numbers of people, buffalo, and animal technicians in a district were associated with an elevated risk of a VDC reporting $\geq 1$ FMD case. The knowledge of high-risk areas and factors associated with the risk of FMD in Nepal could be applied in future disease control programs.

Keywords Foot-and-mouth disease · Nepal · Spatial analysis · Bayesian regression model

Introduction

Foot-and-mouth disease (FMD) is a highly contagious viral disease of domestic and wild cloven-hoofed ungulates and is considered one of the most important diseases affecting animal agriculture production worldwide (Grubman and Baxt 2004). FMD causes significant economic losses mainly due to loss in milk and meat production, mortality of neonates, loss of draft animals due to lameness, and restrictions on trade (James and Rushton 2002; Perry et al. 2002). As in many developing countries, FMD is endemic in Nepal, where the disease has been reported in all 75 districts of the country and confirmed by laboratory testing in 61 districts (Gongal and Karki 2000). Since 1965, four of the seven serotypes of the virus (O, A, C, and Asia 1) have been isolated from cases of FMD in Nepal (Ferris et al. 1992). Economic losses due to FMD infection in Nepal, which has an annual per capita income of $270 (World Bank 2005) and where 92.9% of the human population is primarily involved in agriculture activities (Central Bureau of Statistics of Nepal 2004), have been estimated to be about USD 5.36 million per year (Gongal and Karki 2000).

A study of the economic impact of livestock diseases in rural areas of Nepal estimated that FMD could account for 26% of the overall economic losses in livestock production (Lohani and Rasali 1992).

The epidemiological factors contributing to occurrence of FMD in Nepal are largely unknown. Movement of livestock, which is required to meet the demand for animal draft power and for meat and milk within the country, as well as from India, Bangladesh, and China, has been proposed as a factor promoting reemergence of FMD in Nepal (Ferris et al 1992). FMD prevalence in Nepal has been thought to be higher in areas with a high density of animals and where the human population density also is high (Ferris et al. 1992; Central Bureau of Statistics of Nepal 2004). Animal markets and small ruminant presence also are believed to contribute to FMD spread and transmission (Ferris et al. 1992). The control of FMD in...
Nepal has been difficult because of the limited government resources, lack of trained manpower, poor infrastructure, unstable government, and government policies, such as prohibition of cattle slaughter and lack of access to quality vaccines, as noted elsewhere (Ferris et al. 1992). An understanding of factors associated with FMD risk would help in developing detection, surveillance, and/or control strategies, as well as to inform studies to gain insight into how some factors might be influencing FMD transmission. One approach to uncovering information about risk of disease is to examine whether the distribution of disease coincides with the distribution of hypothesized risk factors.

In general, infectious diseases tend to be distributed in spatial patterns associated in part with the physical and biological factors that influence risk of disease occurrence (Ostfeld et al. 2005). Disease mapping and spatial analysis to explore the nature of such spatial distributions has been widely used in risk modeling to identify high-risk areas or geographically-related risk areas. An objective of such efforts is to identify high risk areas that could subsequently be targeted for control, eradication, and preventive action, as necessary to slow the spread of FMD.

A better understanding of the nature and extent of the contribution of animal, human, and environmental factors to the geographic distribution of FMD in Nepal will help to improve effectiveness and efficiency in the allocation and use of financial and human resources for FMD prevention and control. For example, control and surveillance strategies may be focused on high-risk areas and on factors that contributed to that high risk. Alternatively, regions of low FMD prevalence could be selectively protected to keep incidence low and to prevent introduction of new virus strains. This paper presents the results of spatial analysis and Bayesian modeling to quantify the relation between hypothesized epidemiological factors, or surrogates for epidemiological factors, and the spatial distribution of FMD in Nepal.

Materials and methods

Case definition and data collection

The smallest spatial units at which FMD was reported in Nepal was the Village Development Committees (VDCs), which are administrative units within the district level and are based on geographical and political demarcations, as defined broadly by the Office International des Epizooties guidelines (OIE 2006). The unit of analysis in this study was the district \((n=75)\) and the event of interest was the number of VDCs in the district that reported at least one FMD outbreak in 2004, which was considered to be an approximation for FMD risk per district.

Information on the number of VDCs reporting one or more FMD cases in 2004 and on the values of hypothesized risk factors was obtained from results of a questionnaire administered to each of the 75 district veterinary officers. Data on livestock and human demographics were obtained from the Central Bureau of Statistics’ annual publication (Central Bureau of Statistics of Nepal 2004).

Cluster analysis

The spatial scan statistic (Kulldorff and Nagarwalla 1995), which was implemented using the SatScan software (http://www.satscan.org/), was used to estimate if a geographical aggregation of VDC-reported FMD cases existed for districts in 2004. In the application here, the population at risk was the total number of VDCs per district and cases were the number of VDCs that reported at least one FMD outbreak in the district during 2004. Cases and the population of VDCs at risk of having an FMD case were assumed to be located at the centroid of each district. Circular windows were alternatively centered on each districts’ centroid so that circular windows could be moved over the geographic area. Shapes other than circles, such as ellipsoids, might be used to construct the windows. However, algorithms for quantification of clustering are computationally more intensive for shapes other than circles. Because the objective here was to estimate whether or not observations were spatially independent so that spatial dependence could be adjusted for in the regression model, a circular shape was considered sufficient. The radius of each window was set to vary from zero to 50% of the population at risk, because exploring for high risk of clustering in areas that include >50% of the population at risk is equivalent to explore for areas at low risk outside the cluster, and vice versa (Kulldorff et al. 1998). Significant \((P<0.05)\) clusters were identified using Monte Carlo simulation (Kulldorff and Nagarwalla 1995) and mapped using commercial software (ArcGIS v. 9.2, ESRI Inc. Redlands, CA, USA).

Bayesian model

A Bayesian mixed effects Poisson regression model (Besag et al. 1991) was used to model the association between the number of FMD-positive VDCs in a given district and factors hypothesized to be associated with the risk of having an FMD-positive VDC in a given district. The observed number of FMD-positive VDCs \((O_i)\) within each of the \(i=1, \ldots, 75\) districts was assumed to follow an independent Poisson distribution, with the mean number of cases, or FMD-positive VDCs per district \(\mu_i\), equal to the product of the expected number of cases, or FMD-positive VDCs, in each district \((E_i)\), and the unknown
district-specific relative risk for FMD ($\psi_i$). The estimate of the relative risk parameter $\psi_i$ was a function of $k$ covariates, $x_{ki}$, and possible over-dispersion and spatial autocorrelation were addressed by including two random effects terms, $U_i$ and $S_i$, which represented spatially unstructured and spatially structured effects, respectively:

$$
\log \mu_i = \log E_i + (\beta_0 + \beta_1 x_{1i} + \beta_2 x_{2i} + \ldots + \beta_k x_{ki}) + S_i + U_i
$$

where $x_{1i}$ to $x_{ki}$ denoted 1-to-$k$ factors for districts $i$ modeled as fixed effects with regression coefficients $\beta_1$ to $\beta_k$, respectively.

A preliminary list of 25 hypothesized risk factors ($k$) was examined. The value of the risk factors was transformed to centered values to reduce correlation between beta coefficients and to standardize results. Because the number of risk factors relative to the number of observations was quite large, variables highly correlated with $O_i/E_i$ were screened out using the Spearman correlation test for further testing in the multivariable model. Only covariates found to be correlated at $\alpha<0.10$ with $O_i/E_i$ were chosen to be included as candidate fix effects in the multivariate model (Table 1). Variables for factors related to the district-specific population at risk and species demographics were the number of susceptible buffalo, sheep, goats, pigs, cattle, and yaks present in a district. Variables for factors related to the district-specific population at risk and to species demographics were expressed also as densities and were defined as the number of respective animals per square kilometer. Because animal movement is believed by some to be an important factor for continued reemergence of FMD in Nepal (Ferris et al. 1992), we included district level covariates considered as proxies for animal movement, namely, road density, computed as the ratio of the total length (kilometer) of roads in a district to the area (square kilometer) of the district, the number of live animal markets in the district, the number of animals slaughtered daily for meat supply in the district, and the number of live animal imports and exports. Because human demographics are believed to be associated with risks of FMD and other diseases (Harrington et al. 2005; Rivas et al. 2006), variables related to human demographics or to socio-economic attributes of a district included total human population, human population density, proportion of the population that was urban, which was defined as the proportion of people who lived in the urban areas of the district and who did not raise food animals for their livelihood, and literacy rate, which was expressed as the percentage of the district population that completed 5 years of schooling. As suggested elsewhere, Village Animal Health Workers (VAHWs) have been instrumental in surveillance, reporting, prevention, and treatment of animal diseases in other countries (Leyland and Catley 2002). Variables related to the ability of the animal health service to diagnose, report, and prevent FMD spread included the number of veterinarians, the number of technicians, the number of VAHWs per district, and the proportion of the population having access to veterinary care in the district. Two other variables were included to indicate whether a district shared a border with one of the two neighboring countries, as indicated by: China (yes/no) or India (yes/no), to account for disease spread from these countries.

The spatially unstructured random effect, $U_i$, accounted for unmeasured explanatory variables that were distributed randomly between areas, whereas the structured random effect, $S_i$, accounted for unmeasured explanatory variables that had some spatial distribution. A non-informative Gaussian prior with mean equal to zero and precision equal to 0.0001 was assumed for the intercept, $\alpha$, and the regression coefficients, $\beta$, were assumed to reflect a lack of prior knowledge about the strength of any association between FMD risk and the covariates. Prior information about $U_i$, was specified as $U_i \sim N(0, \tau_u)$, where $\tau_u = 1/\sigma_u^2$. The random effect, $S_i$, was assigned an intrinsic Gaussian autoregressive (CAR) structure whereby the prior distribution of each $S_i$ was conditional on the value of $S_j$. The mean value for $S_i$ was a weighted average of the neighboring structured random effects, where the variance, $\sigma_s^2$, controlled the strength of the local spatial dependence (Clayton and Kaldor 1987; Besag et al. 1991; Souza et al. 2007). The CAR prior structure allowed for smoothing of the relative risk in district $i$, conditional on the risks in all other districts $i\neq j$. Neighbor specification was based on adjacency of districts such that the value for each district was influenced only by districts that share a common border with it (Benardinelli et al. 1995; Best et al. 1999). This was considered a conservative simplification of the model because the influence that districts have on their neighbors is likely dictated by factors that are difficult to measure in Nepal, such as intensity of formal and animal and human movements, similarity of cultural and religious practices, and other social and economic factors. The standard deviations of $S$ and $U$ were used to compare the variability of structured and unstructured random effects and to allow inferences as to whether risk factors not included in the model were spatially distributed or not. Model computations were done using WinBUGS (Spiegelhalter et al. 1999), with 100,000 iterations, and the first 500 ‘burn-in’ samples were discarded. Several combinations of risk factors were evaluated to identify the combination that best predicted the $O/E_i$, as indicated by the value of the deviance information criterion (DIC) value (Spiegelhalter et al. 2002) Alternative combinations with variables biologically associated to those included in the final model, irrespectively of their Spearman’s rank.
Table 1 Spearman’s rank correlation coefficient of the association between the ratio of observed and expected FMD cases (Oi/Ei), and factors hypothesized to be associated with the risk for FMD in Nepal in 2004. The posterior median of the risk ratios of covariates included in five candidate regression models examined during model fitting process is also indicated. The models included structured and unstructured random effect. Model 1 provided the best fit, as indicated by the lowest value of the Deviance Information Criterion.

| Covariates                          | Spearman’s rho value | P value | Risk ratios (95% probability intervals) |
|-------------------------------------|----------------------|---------|----------------------------------------|
|                                     |                      |         | Model 1                                 | Model 2                    | Model 3     | Model 4     | Model 5     |
| Population of humans                | 0.456                | <0.01*  | 1.14 (0.8, 1.7)                         | -                         | 1.15 (0.7, 1.7) | -          | 1.36 (0.9, 1.9) |
| Human population density (persons/km²) | 0.433                | <0.01*  | -                                      | 1.33 (1.0, 1.7)           | -          | -          | -           |
| Number of buffaloes                 | 0.429                | <0.01*  | 1.76 (1.2, 2.6)                         | -                         | 1.79 (1.2, 2.6) | 1.85 (1.3, 2.7) | 1.65 (1.1, 2.5) |
| Number of technicians               | 0.426                | <0.01*  | 1.31 (1.0, 1.8)                         | 1.36 (1.0, 1.8)           | 1.23 (0.9, 1.7) | 1.25 (0.9, 1.7) | 1.29 (0.9, 1.8) |
| Number of animals slaughtered daily | 0.416                | <0.01*  | -                                      | -                         | -          | -          | -           |
| Urban population (%)                | 0.369                | <0.01*  | -                                      | -                         | -          | -          | -           |
| Buffalo density (buffaloes/km²)     | 0.355                | <0.01*  | -                                      | 1.41 (1.0, 1.9)           | -          | -          | -           |
| Number of veterinarians             | 0.291                | 0.01*   | -                                      | -                         | -          | -          | -           |
| Number of village animal health workers | 0.272                | 0.02*   | -                                      | -                         | -          | -          | -           |
| Cattle density (cattle/km²)         | 0.272                | 0.02*   | -                                      | -                         | -          | -          | -           |
| Status of border with China (yes, no) | -0.263               | 0.02*   | -                                      | -                         | -          | -          | -           |
| Number of cattle                    | 0.242                | 0.04*   | -                                      | -                         | -          | -          | -           |
| Road density (km road/km²)          | 0.237                | 0.04*   | -                                      | -                         | 1.03 (0.6, 1.7) | 1.02 (1.0, 1.6) | -           |
| Goat density (goats/km²)            | 0.205                | 0.08*   | -                                      | -                         | -          | -          | -           |
| Number of goats                     | 0.186                | 0.11    | -                                      | -                         | -          | -          | -           |
| Number of pigs                      | -0.153               | 0.19    | -                                      | -                         | -          | -          | -           |
| Literacy rate (%)                  | 0.139                | 0.23    | -                                      | -                         | -          | 1.06 (0.8, 1.4) | -           |
| Number of live animal markets      | -0.121               | 0.30    | -                                      | -                         | -          | -          | -           |
| Number of live animals exported     | -0.107               | 0.36    | -                                      | -                         | -          | -          | -           |
| Number of animal imports           | -0.107               | 0.36    | -                                      | -                         | -          | -          | -           |
| Pig density (pigs/km²)              | -0.084               | 0.47    | -                                      | -                         | -          | -          | -           |
| Number of live animals imported     | -0.074               | 0.53    | -                                      | -                         | -          | -          | -           |
| Number of sheep                    | -0.059               | 0.61    | -                                      | -                         | -          | -          | -           |
| Status of border with India (yes, no) | 0.051                | 0.66    | -                                      | -                         | -          | -          | -           |
| Sheep density (sheep/km²)           | 0.042                | 0.72    | -                                      | -                         | -          | -          | -           |
correlation coefficient value, were also fitted to explore the nature of the relations among factors influencing FMD risk. Hyper prior distributions were assigned to the precision terms of the random effects $\tau_u = 1/\sigma_u^2 \sim Gamma(0.5, 0.0005)$ and $\tau_s = 1/\sigma_s^2 \sim Gamma(0.5, 0.0005)$. Non-informative priors and hyper priors used in here were similar to those used by others (Kelsall and Wakefield 1999; Wakefield et al. 2000; Stevenson et al. 2005). The posterior median estimate of $y_i$, denoted as $\hat{y}_i$, was computed to assess the influence of unknown factors on the distribution of FMD relative risk in different districts and mapped using ArcGIS v. 9.2 (Arc-Map™ 9.0., ESRI Inc. Redlands, CA, USA).

Bayesian model

The $O/E_i$ ratios estimated from the model and plotted as choropleth maps indicated considerable variability in risk of VDC-reported FMD cases throughout the country (Fig. 1b). There was some visual evidence of an aggregation of districts with elevated risks ($O/E_i > 1$) in the terai and hill regions of central Nepal. High risk districts also were found in the same areas as identified by cluster analysis namely, Kathmandu, Nuwakot, Mahottari, and Sarlahi.

Based on results of the Spearman’s rank correlation, 14 covariates were highly correlated ($P<0.1$) with the $O/E_i$ ratio at the district level (Table 1). After considering these 14 candidate variables, district-specific variables in the model providing the best fit (DIC=262.29) included the total number of people, the total number of buffalo, the total number of technicians, the interaction between the number of buffalo and the number of technicians, and random effects (model 1: Table 1). Irrespective of geographical region, in districts where the buffalo population was greater than the median of all districts, an increase in the number of technicians was associated with a decrease in the predicted number of VDCs reporting $\geq 1$ cases of FMD (Fig. 3). The exponential of the posterior median value of the regression coefficients suggested that each standardized unit increase in the values of the number of people, buffaloes, and technicians was associated with an increase in the risk of a VDC reporting $\geq 1$ cases of FMD in the district by factors of 1.14, 1.76, and 1.31, respectively (Table 1).

With the set of non-informative priors specified for the regression coefficients in the model, only number of buffaloes and an interaction between number of buffaloes and technicians was found to be statistically significant. Inclusion of covariates representing population densities, such as human and animal densities, gave a poorer fit than models with the same covariates expressed as absolute numbers. In model 1, the risk associated with the size of the

---

**Table 1** (continued)

| Covariates | Spearman’s rho value | $P$ value | Risk ratios (95% probability intervals) |
|------------|----------------------|-----------|----------------------------------------|
| Interaction+ | -                    | -         | 0.58$^b$ (0.4, 0.8) 0.76$^b$ (0.5, 1.0) 0.65$^a$ (0.4, 0.9) 0.65$^a$ (0.4, 0.9) 0.62$^a$ (0.4, 0.9) |
| Deviance Information Criterion | -                    | -         | 262.29                                 |

- not applicable; *correlation at $\alpha<0.10$; + interaction between number of technicians and (a) number of buffaloes and (b) buffalo density
human population in a district increased slightly with inclusion of the number of animals slaughtered daily and of the road density (Table 1). The risk associated with the number of animals slaughtered daily in a district increased when the human population variable was removed from the model. Further, risk associated with the size of the human population increased slightly when literacy was included in the model (Table 1).

The district-specific posterior estimates of FMD risk ($\hat{\Psi}_i$), or the relative risk obtained after adjusting for effects of the three putative risk factors and accounting for structured and unstructured random effects, ranged from 0.1 to 8.3 (Fig. 4a). There was a 25-fold difference in the values estimated for the spatially structured (SDs=0.036) and for the unstructured random effects (SDu=0.884). As shown by the choropleth map, the exponential of the posterior medians of unstructured random effects, which represented relative risks associated with unknown factors that were not explicitly or implicitly addressed in the model, did not appear visually to have a clear discernible geographic pattern (Fig. 4b).

**Discussion**

The present study identified districts in Nepal where the risk of FMD, as estimated by FMD reports from the VDCs, was significantly higher than that for the average district and it identified factors associated with the variation in risk among the districts. As indicated by results of the multivariate Bayesian model, district factors associated with a high risk of FMD were a high human population, a high buffalo population, and a large number of animal health technicians. These factors accounted for most of the observed clustering, as indicated by the higher value of SDu (0.884), compared with the value of the SDs (0.036).
The finding that districts with many buffalo and technicians had higher relative risks of VDC-reported FMD (1.76, PI: 1.24, 2.62; and 1.31, PI: 0.96, 1.84, respectively), compared with districts with fewer buffalo and technicians, suggests possible links between FMD risk and buffaloes and technical support. Risk was found to vary differently depending on the number of technicians and buffaloes in the district, as indicated by the significant interaction between number of buffalo and number of technicians (RR: 0.58, PI: 0.38, 0.83). Alternative explanations could account for the interaction found between number of buffalo and number of technicians and its association with FMD risk. The sensitivity of reporting FMD could have been higher in districts with a large number of technicians or, perhaps, technicians in some way contributed to transmission of the disease (Figs. 2 and 3). A high risk associated with a larger number of buffaloes could relate to enhanced informal owner surveillance that might be expected in areas where buffaloes are raised. Buffaloes are considerably more expensive than other cattle in Nepal and other countries in South Asia and, as such, generally would be expected to command greater attention to their health and productivity. Consequently, an observed higher risk of FMD in regions with a large number of buffaloes could represent greater awareness and scrutiny for disease, and thus more rigorous observation and reporting of suspicious FMD cases, which may have biased the results presented here. Alternatively, the findings could be interpreted to indicate that buffalo serve in some biological way to increase risk of FMD, perhaps through a virus-host relationship that would foster transmission. Although water buffaloes (*Bubalus bubalis*) can effectively transmit the FMD virus to other susceptible species (Gomes et al. 1997; Dutta et al. 1983), and the carrier state has been found in Indian buffaloes (Samara and Pinto 1983; Barros et al. 2006), there are no reports suggesting that water buffaloes are more likely to become infected or that they are more likely to shed large amounts of viruses. If buffalo were more susceptible or if they did shed more virus for longer periods, one would expect districts with many buffaloes to have greater risks of FMD, compared with districts with fewer buffaloes.

The finding of a positive association (RR: 1.14, PI: 0.8-1.7) between a district’s human population and the number of VDCs reporting FMD suggests that the size of...
the human population in a district, or some unmeasured factors associated with the size of the human population, was in some way associated with increasing the risk of FMD. Features of human demographics, such as population density, socio-economic status, and literacy rate, have been discussed elsewhere as possibly having a role in fostering animal diseases, including FMD (Rivas et al. 2006; Harrington et al. 2005), where human population size might be considered as a surrogate for the extent of animal movement expected in a region (Rivas et al. 2006). In the model presented here, human population size was used in the model as a proxy for several factors thought to contribute to FMD spread, including transmission of FMD by people and animal movement, which would be expected to increase in response to higher demands for meat and milk in densely populated areas. Consideration of candidate models, which had very small differences in DIC, was helpful in assessing use of the human population variable as a fixed effect. These candidate models were compared with models that considered road density and number of animals slaughtered, which were highly correlated with size of the human population and could also be considered as a proxy for animal movement. The positive association found between the number of animals slaughtered daily in a district and the FMD risk when human population was removed from the model, although not statistically significant, (Table 1) suggested that human population variable could be acting as a proxy for the number of animals slaughtered daily. An important consideration is that, apart from accounting for factors that have a positive effect on FMD, inclusion of a human population variable as a fixed effect also would identify other correlated factors that were negatively associated with FMD risk, or that could have equivocal effects. For example, as reported elsewhere, higher levels of education could be expected to help reduce FMD risk but also could increase sensitivity of reporting (Saini et al. 1992), thereby presenting as little or no effect either positive or negative. Here, inclusion of the literacy rate variable into the model was associated with an apparent increased risk of FMD attributed to human population size, which might suggest that for districts with a population that is more aware and educated, the sensitivity of reporting FMD would be higher because FMD cases would more likely be reported than in districts with a lesser literacy rate (Table 1).

The 25-fold difference in standard deviations of the structured (SDs=0.036) and unstructured (SDu=0.884) random effects suggests that most of the variation in FMD risk that was not explained by factors included in the model was associated with an apparent increased risk of FMD attributed to human population size, which might suggest that for districts with a population that is more aware and educated, the sensitivity of reporting FMD would be higher because FMD cases would more likely be reported than in districts with a lesser literacy rate (Table 1).

In summary, results of this study indicate that the FMD risk in Nepal in 2004 was spatially clustered in specific areas of the country, and that the elevated risk in those areas was largely associated with the size of human and buffalo populations and with the number of technicians. Further studies should be directed to clarify the reasons for these associations, in order to increase the knowledge on the epidemiological dynamics of FMD in Nepal and to improve the efficiency of resource allocation and control efforts in the country.

Acknowledgments The project was funded in part by the National Center for Medical Intelligence. We gratefully acknowledge the assistance of Dr. Raju Gautam in the administration of the questionnaires and the collaboration of the district veterinarians in Nepal who participated in the survey.
