Spatial epidemiology of bovine tuberculosis in Mexico

Horacio Zendejas Martínez, Feliciano Milián Suazo, José Quintín Cuador Gil, Gustavo Cruz Bello, Ana María Anaya Escalera, Gabriel Huitrón Márquez & Leticia García Casanova

Summary

The purpose of this study was to use geographic information systems (GIS) and geo-statistical methods of ordinary kriging to predict the prevalence and distribution of bovine tuberculosis (TB) in Jalisco, Mexico. A random sample of 2 287 herds selected from a set of 48 766 was used for the analysis. Spatial location of herds was obtained by either a personal global positioning system (GPS), a database from the *Instituto Nacional de Estadística Geografía e Informática* (INEGI) or Google® Earth. Information on TB prevalence was provided by the Jalisco Commission for the Control and Eradication of Tuberculosis (COEETB). Prediction of TB was obtained using ordinary kriging in the geostatistical analyst module in ArcView™ 8. A predicted high prevalence area of TB matching the distribution of dairy cattle was observed. This prediction was in agreement with the prevalence calculated on the total 48 766 herds. Validation was performed taking estimated values of TB prevalence at each municipality, extracted from the kriging surface and then compared with the real prevalence values using a correlation test, giving a value of 0.78, indicating that GIS and kriging are reliable tools for the estimation of TB distribution based on a random sample. This resulted in a significant savings of resources.

Keywords

Cattle, Geographic information system, Kriging, Mexico, *Mycobacterium bovis*, Tuberculosis.

Epidemiologia spaziale applicata alla tubercolosi bovina in Messico

Riassunto

Scopo di questo lavoro è stato l'utilizzo dei sistemi informativi geografici e dei metodi di "ordinary kriging" per prevedere la prevalenza e la distribuzione della tubercolosi bovina (TB) a Jalisco, Messico. È stato selezionato un campione casuale di 2.287 capi da una popolazione di 48.766. La localizzazione spaziale dei capi è stata ottenuto per mezzo di un GPS o da un database de l' Instituto Nacional de Estadística Geografía e Informática (INEGI) o attraverso l'uso di Google® Earth. Le informazioni sulla prevalenza della TB sono state fornite dalla Commissione Jalisco per il controllo e l'eradicazione della tubercolosi (COEETB). La previsione per la tubercolosi bovina è stata ottenuta utilizzando kriging nel modulo di analisi geostatistica in ArcView™ 8. Il sistema ha predetto alta prevalenza di tubercolosi bovina coincidente con la distribuzione degli allevamenti. I livelli di prevalenza predetti dal modello sono risultati in accordo con i livelli di prevalenza calcolati sul totale di 48.766 capi. La validazione è stata effettuata considerando i valori stimati di prevalenza di TB in ogni municipalità, estratti dalla superficie kriging e quindi comparati con i valori reali di prevalenza utilizzando test di correlazione, assegnando un valore di 0.78, indicante che il GIS e il kriging sono strumenti affidabili per la stima, basata sull'analisi di un campione casuale, della distribuzione della TB, ne consegue un notevole risparmio di risorse.

Parole chiave

Bovini, Kriging, *Mycobacterium bovis*, Messico, Sistema informativo geografico, Tubercolosi.

Programa Nacional de Epidemiología, Instituto Nacional de Investigaciones Forestales, Agrícolas y Pecuarias, Apartado Postal 1-1168, Querétaro, Qro. 7601, Mexico
mchoraciozendejas@gmail.com

© IZS A&M 2007

Introduction

Like many countries, Mexico has experienced an increase in the incidence of tuberculosis (TB) in humans (1). The increase in population density in suburban areas and poverty (8), the presence of the human immunodeficiency virus (HIV) and some other debilitating diseases (4) and the presence of strains resistant to antibiotics, are some of the factors responsible for this phenomenon.

The primary aetiology of TB in humans is *Mycobacterium tuberculosis*, however, *M. bovis*, the agent of TB in cattle is responsible for about 5% to 10% of all human cases (13). Infections of humans from cattle are caused by close contact with infected animals or by the ingestion of unpasteurised milk or cheese made with raw milk and infected secretions (6, 7). Thirty percent of the 7 million litres of milk produced annually in Mexico do not go to pasteurising plants, which is one of the reasons why TB has to be eliminated in cattle.

TB in cattle also needs to be eliminated for economic reasons. Mexico exports an average of 1.2 million calves every year to the United States and Mexico has been regionalised in accordance with disease prevalence, i.e. only free states are allowed to export animals. This constraint represents a problem for states that do not have accreditation of disease-free status as animal movements are banned, not only to foreign countries, but also to other states within the country. For non-accredited states to be included in the exporting programme, it is necessary to provide scientific evidence of TB status. Consequently, these states need to estimate prevalence rapidly at low cost and should provide scientifically valid data that will be acceptable to trade partners. We believe this can be accomplished by using geographic information systems (GIS) and geo-statistical methods in a representative sample of herds.

GIS is one of the tools for organising and analysing information for a better understanding of disease epidemiology. GIS consists of computational systems, databases and digitalised spatial location (points, lines, polygons). GIS has been widely used to study the spatial and temporal distribution of diseases in different parts of the world (2, 3, 16).

In veterinary epidemiology, the need for mapping locations of herds is a priority. In any epidemic disease it helps to monitor and analyse disease spread and it also provides a very useful tool to evaluate different strategies of control measures. GIS can be incorporated into a more general information system aimed at recording and reporting information but, at the same time, specific functions and tools can be integrated for cluster analysis, spatio-temporal trends and modelling of disease distribution (12). Perez *et al.* (14) applied spatial statistics and monitoring data to identify clusters of bovine TB in Argentina, with the aim of focusing efforts on the eradication of the disease. GIS has also been used to determine the association between TB status and distances to badger setts (10) to decide on control strategies to solve the problem.

GIS frequently interacts with specific and advanced spatial methodologies to accomplish the objectives set. Geo-statistics is one example through which it is possible to estimate unknown values on a continuous surface for a specific variable using sampled point values (5, 15), minimising the variance of error in the estimation (11). Geo-statistics can be applied to predict disease prevalence in areas that have not been sampled, starting from point prevalence observations. In this study, prevalence and distribution estimates of bovine TB in Mexico have been calculated through geo-statistical techniques, such as ordinary kriging.

Materials and methods

Obtaining field observations can be expensive, laborious and time consuming. Thus, it can be extremely useful to predict a location value where we have no recorded field observation. The process of obtaining such values for a variable of interest at an unsampled location based on neighbour measurements is called interpolation. Kriging is a group of geo-statistical techniques used to interpolate the values at unobserved locations from

© IZS A&M 2007

observations of its value at nearby locations. It returns unbiased spatial predictions (the mean of errors is 0) and the prediction mean-squared error, a measure of uncertainty or variability in the predicted values is minimised. Kriging uses the semi-variogram, a function of distance and direction separating two locations, to analyse and quantify the spatial autocorrelation in the data. The semi-variogram is then used to establish the contribution of each data point to the prediction of new values at the unsampled locations (8). The most important statistical assumption supporting kriging is stationarity, which means that statistical characteristics (such as mean and variance) do not depend on the exact spatial locations. Thus, the mean and variance of a variable at one location is equal to the mean and variance at another location. Furthermore, the correlation between any two measured points depends only on the vector that separates them and not on their exact location. When the working data cannot satisfy this assumption of stationarity, there are several techniques that can be used to perform the kriging method (e.g. detrending) (4, 9, 11, 17).

To emphasise the usefulness of GIS techniques in predicting the prevalence of TB, one state of Mexico was selected as a study area. Information collected between 1995 and 2005 regarding tuberculin-testing was analysed. A complete database of 2 400 000 records from 48 776 herds was accessed. Bovine TB prevalence was calculated for each heard by dividing the number of positive animals to the tuberculin test by the total number of animals tested, multiplied by 100. A positive bovine TB herd was one with herd-prevalence greater than zero.

A total of 2 286 herds, a little more than the 2 074 estimated with 6% population prevalence, 1% error and a 95% confidence level, were randomly selected from a total of 48 766 to be used in the study. The sample was stratified according to cattle population (per county) and breed (beef or dairy). All selected herds were geo-referenced using a global positioning system (GPS) (Geko 301, Garmin Ltd), data from the National Institute for Statistics, Geography and Informatics (INEGI) or by Google® Earth.

A prediction of bovine TB prevalence was calculated using geo-statistical techniques (ordinary kriging). The spatial pattern of the sampled herds was analysed by the semi-variogram technique using the geo-statistical analyst extension of the ArcView™ 8.1 package (Environmental Systems Research Institute, Inc., Redlands, California). Point prevalence calculated in each sampled herd was interpolated through ordinary kriging and a continuous surface was calculated to estimate the unsampled points.

Results

A total of 2 286 herds in the Mexican state of Jalisco were used in the study.

Figure 1 shows the continuous surface generated by ordinary kriging of TB-prevalence in the study area. Prediction clearly shows free, low and high prevalence areas. As expected, high prevalence areas correspond to high population densities of dairy cattle, where it is known that transmission and persistence of bovine TB is more likely to occur. To evaluate accuracy of prediction, estimated values of TB prevalence at each municipality extracted from the kriging surface were compared with the real prevalence values using a correlation test and a value of 0.78 was obtained. Figure 2 shows the distribution of TB based on the total number of herds, namely: 48 776. This map was generated by using the results of the tuberculin test applied to the 48 766 herds in Jalisco, Mexico, that were provided by the official government TB campaign.

Discussion

It is well known that disease distribution is not random. Rather, it follows a distribution pattern similar to that of the aetiology, host and vectors, when these are involved, and the appropriate environmental conditions; TB is no different. Figures 1 and 2 show the prevalence rates of bovine TB in the study area. Those areas with higher prevalence rates correspond to areas of higher dairy population density, which clearly indicates that dairy cattle are the main risk for TB persistence and dissemination and that, in order to eliminate the disease, dairy herds must be the focus of attention. Nevertheless, it can be presumed that additional information, such as herd size, herd-population dynamics, breed population density, regional population density and regional population dynamics would increase the accuracy of the prediction. Currently, all this type of information is being collected to complete the study. Kriging is an interpolation method that has proved to be useful in numerous practical applications. However, it is necessary that specific assumptions are

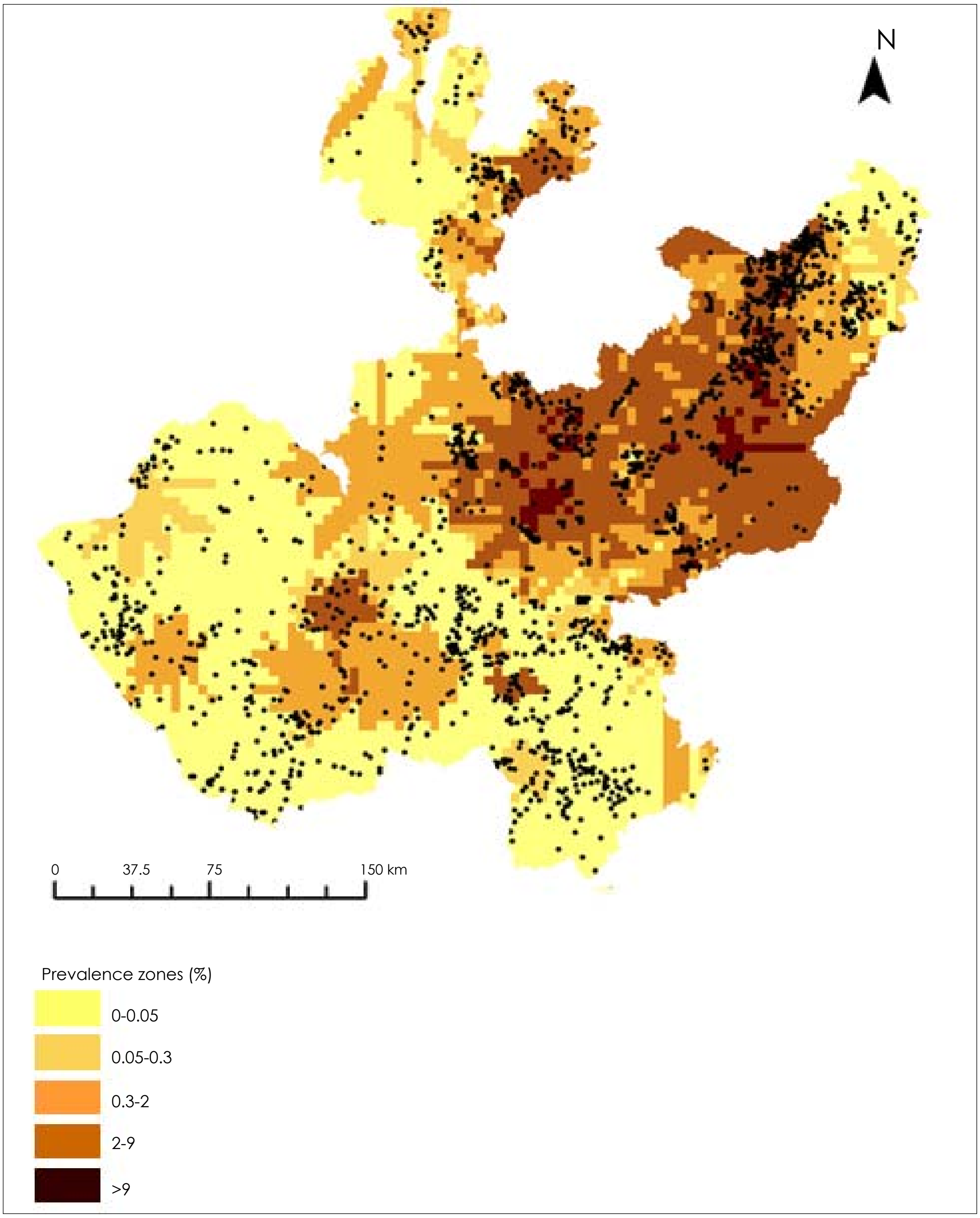

Figure 1
Kriging-predicted bovine tuberculosis prevalence zones in the state of Jalisco, based on a 2 286 sample of herds
Dots show the geographic location of herds included in the sample

© IZS A&M 2007

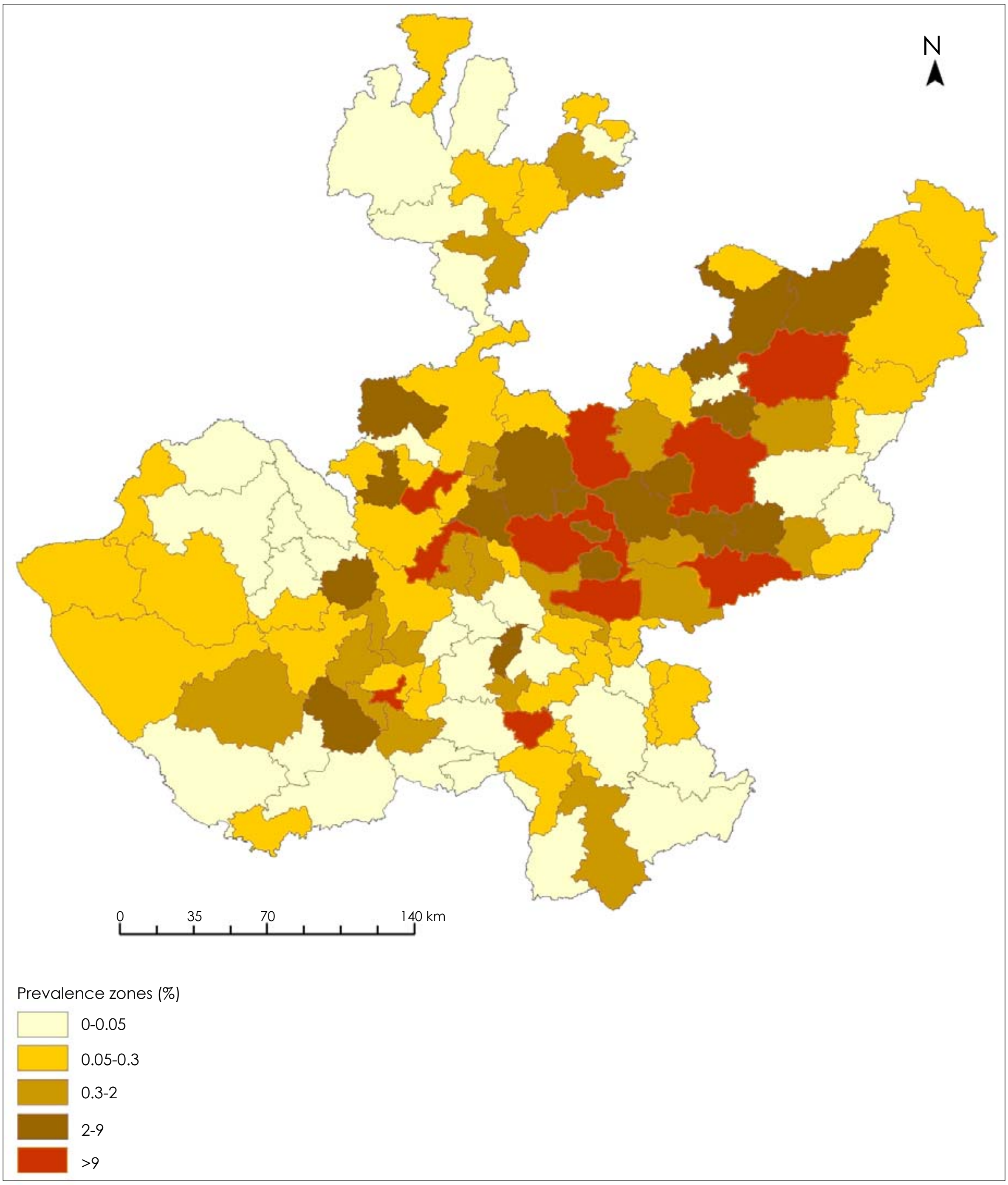

Figure 2
Bovine tuberculosis prevalence zones in the State of Jalisco, by municipality, based on a total of 48 766 herds sampled using the scanning procedure

respected in the dataset analysed. In this case, ordinary kriging was used as a tool to estimate bovine TB prevalence in areas that had not been sampled, taking into account spatial autocorrelation and variability. Further analysis would be helpful to verify the optimal sample size required to obtain reliable estimates in the entire study area, saving resources as far as possible.

Acknowledgments

We would like to thank the Jalisco Commission for the Control and Eradication of Tuberculosis (COEETB) for all tuberculin-testing data. Sincere thanks are also extended to Marina Duran for helping in the analysis of the database in SAS.

Grant support

This study was funded by the *Consejo Nacional de Ciencia y Tecnología de México* (CONACYT) Project No. SAGARPA-2003-C01-128.

References

1. Anon. 1997. Epidemiología. Sistema Nacional de Vigilancia Epidemiológica, Sistema Único de Información, México. *Sist Nac Vigil Epidemiól*, **2**, 1-4.
2. Briggs D.J. & Elliott P. 1995. The use of geographical information systems in studies on environment and health. *World Health Stat Q*, **48**, 85-94.
3. Brooker S. & Michael E. 2000. The potential of geographical information systems and remote sensing in the epidemiology and control of human helminth infections. *Adv Parasitol*, **47**, 246-279.
4. Fernández D. 1997. El control de la tuberculosis, México. *Neumología Cirugía Tórax*, **56**, 116-118.
5. Gotway C.A.1991. Fitting semivariogram models by weighted least squares. Short Note, *Comp Geosci*, **17** (1), 171-172.
6. Grange J.M. 1995. Human aspects of *Mycobacterium bovis* infection. *In Mycobacterium bovis* infection in animals and humans (C.O. Thoen & J.H. Steele, eds). Iowa State University Press, Ames, 29-45.
7. Guillespie J.H. & Timoney J.F. 1983. The *Mycobacterium* genus. *In* Hagan and Bruner's microbiology and infectious diseases of domestic animals, 8[t]h Ed. Cornell University Press, New York, 270-302.
8. Gutierrez M., Vincent B. & Aubert D. 1998. Molecular fingerprinting of Mycobacterium tuberculosis and risk factors for tuberculosis transmission in Paris, France and surrounding area. *J Clin Microbiol*, **32**, 486-492.
9. Krivoruchko K. & Gotwayb C.A. 2004. Creating exposure maps using kriging. *Publ Health GIS News Info*, January, 1-8.
10. Martin S.W., Eves J.A., Dolan L.A., Hammond R.F. Griffin J.M., Collins J.D. & Shoukri M.M. 1997. The association between the bovine tuberculosis status of herds in the east Offaly Project Area, and the distance to badger sett, 1988-1993. *Prev Vet Med*, **31**, 113-125.
11. Matheron G. 1997. La théorie des variables regionalisées et ses applications. Les Cahiers du Centre de Morphologie Mathématique de Fontainebleau. Fascicule 5, École de Mines de Paris, Fontainebleau, 212.
12. Norstrom M. 2001. Geographical information system (GIS) as a tool in surveillance and monitoring of animal diseases. *Acta Vet Scan. Suppl.*, **94**, 79-85.
13. O'Reillly L.M. & Daborn C.J. 1995. The epidemiology of *Mycobacterium bovis* infections en animals and man: a review. *Tubercle Lung Dis*, **6**, 1-46.
14. Perez A.M., Ward M.P., Torres P. & Ritacco V. 2000. Use of spatial statistics and monitoring data to identify clustering of bovine tuberculosis in Argentina. *Prev Vet Med*, 56, 63-74.
15. Rivoirard J. & Guiblin P. 1997. Global estimation variance in presence of conditioning parameters. *In* Geostatistics Wollongong '96, Vol. 1 (E.Y. Baafi & N.A. Schofields, eds). Centre de Géostatistique, École des Mines de Paris, Fontainebleau, 246-257.
16. Robinson T.P. 2000. Spatial statistics and geographical information systems in epidemiology and health public. *Adv Parasitol*, **47**, 82-127.
17. Zhang R., Myers D.E. & Warrick A.W.1992. Estimation of the spatial distribution of soil chemical using pseudo cross-variograms. *Soil Sci Soc Am J*, **56** (5), 1444-1452.

 © IZS A&M 2007