STATISTICS IN MEDICINE
Statist. Med. 2001; **20**:2319–2335 (DOI: 10.1002/sim.844)

Diffusion and prediction of Leishmaniasis in a large metropolitan area in Brazil with a Bayesian space–time model

Renato M. Assunção[1,*,†], Ilka A. Reis[1] and Claudia Di Lorenzo Oliveira[2]

[1] *UFMG, Departamento de Estatística, Caixa Postal 702, Belo Horizonte, Minas Gerais–30123-970, Brazil*
[2] *Núcleo de Vigilância Epidemiologica, Hospital das Clinicas, Belo Horizonte, Minas Gerais, 30130–100, Brazil*

SUMMARY

We present results from an analysis of human visceral Leishmaniasis cases based on public health records of Belo Horizonte, Brazil, from 1994 to 1997. The main emphasis in this study is on the development of a spatial statistical model to map and project the rates of visceral Leishmaniasis in Belo Horizonte. The model allows for space–time interaction and it is based on a hierarchical Bayesian approach. We assume that the underlying rates evolve in time according to a polynomial trend specific to each small area in the region. The parameters of these polynomials receive a spatial distribution in the form of an autonormal distribution. While the raw rates are extremely noisy and inadequate to support decisions, the resulting smoothed rates estimates are considerably less affected by small area issues and provide very clear directions to implement public health actions. Copyright © 2001 John Wiley & Sons, Ltd.

1. INTRODUCTION

Visceral Leishmaniasis is a typically rural zoonotic disease of wide prevalence in many countries from Europe, Asia and the Americas. Its cycle includes man, the fox (*Lycalopex vetulus*) and the dog (*Canis familiaris*) as wild and domestic reservoirs, respectively, and the sandfly *Lutzomya longipalpis* as a vector. The importance of the dog as a reservoir is well known [1–3]. Recently, several urban foci have been described in large Brazilian cities such as São Luís [4], Natal [5], Teresina [6, 7] and Belo Horizonte [8].

In March 1989, one case of human visceral Leishmaniasis (HVL) was diagnosed in Sabará, a municipality located in the Belo Horizonte metropolitan region [8]. The Universidade Federal de Minas Gerais (UFMG) and Fundação Nacional de Saúde (FNS) worked together to detect new human cases and undertake control actions in that region.

*Correspondence to: Renato M. Assunção, UFMG, Departamento de Estatística, Caixa Postal 702, Belo Horizonte, Minas Gerais, 30123-970, Brazil
†E-mail: assuncao@est.ufmg.br

Contract/grant sponsor: FAPEMIG
Contract/grant sponsor: CNPq
Contract/grant sponsor: FUNDEP

Received July 1999
Accepted August 2000

In 1992, seropositive dogs were detected in the Belo Horizonte Northeast region. In 1993, 2288 dog blood samples were screened and 5.8 per cent tested positive. One year later, 2951 dogs were screened and 7.6 per cent tested positive. In 1994, there were 3.1 per cent seropositive dogs among 12 317 examined. In 1995, 107 850 dogs were analysed and 2.9 per cent tested positive. These numbers changed in 1996 to 110 000 samples collected and 3.8 per cent testing positive. The Brazilian Health Ministry recommends the elimination of all positive dogs and indoor DDT nebulization and these recommendations were followed since the beginning of the epidemics.

The canine prevalence has always preceded the occurrence of human cases. In 1994, 29 human cases were diagnosed in Belo Horizonte, all located in the Eastern and Northeastern regions [9]. In 1995, 46 human cases were reported to the public health officials. In 1996, 45 human cases were detected and, only in the first semester of 1997, 40 cases were reported. Among the 160 human cases reported from 1994 to 1997, there were 17 deaths, implying a lethality of 10 per cent.

In spite of the control actions undertaken, the number of new human cases did not decrease substantially over time as expected by public health officials. The health control programme recommended by the Brazilian Health Ministry is an expensive programme amounting to approximately U.S. $ 1.5 million in 1995 in Belo Horizonte. Considering the high cost involved, it would be necessary to rank the areas of the city according to a priority score. Based on that score, it would then be possible to aim for a more focused action of the municipality health officials, accomplishing a decrease in costs and providing better service. Since the small number cases in each area produce very unstable rates, this brings the need for a better method to estimate the underlying risks.

The objective of this work is to map the rates of HLV in Belo Horizonte, to estimate their spatial risk structure and to follow their evolution from 1994 to 1997. Additionally, we construct a model to project the rates providing subsidies for public health interventions in the higher risk areas. We use a space–time Bayesian model proposed by Bernadinelli *et al.* [10] but with a slightly different parameterization. In contrast to their results, we find strong evidence of space–time interaction in our data. The model filters out the enormous random variation of the rates which is unrelated to the underlying risk. We also use the parametric model to project the rates for the next years and use a cross-validation type of procedure to select an appropriate model for the rates. The model brings great benefits to orient public health actions since an extremely noisy data set can be smoothed out to produce plots and maps where the information on the disease dynamics is evident, at the cost of weak constraints on the spatial smoothness of the disease rates.

2. THE DATA

Belo Horizonte is a large Brazilian city with 2091 thousand inhabitants and 600 thousand households according to the last 1995 census. Located in the Southeast part of Brazil, it ranks third in size in the country and it has a large service and industrial labour population. Starting in 1960, it had a fast population increase stabilizing between 1990 and 1995.

The Belo Horizonte municipality is partitioned into nine public health administrative regions, called Sanitary Districts (SD). Each SD is further partitioned in *áreas de abrangência*, here-

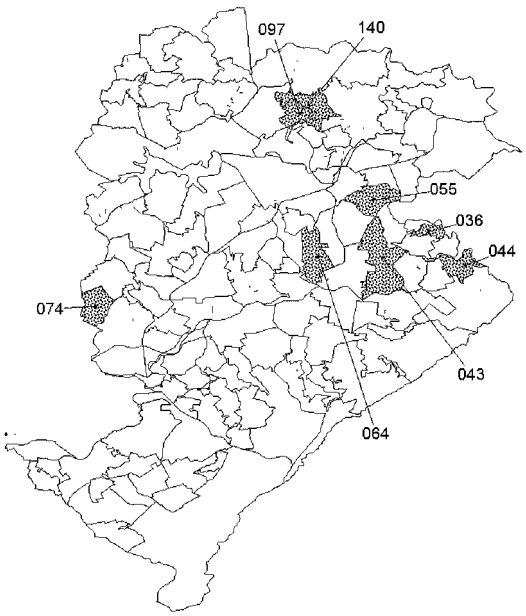

Figure 1. Map of Belo Horizonte divided into 117 health zones.

after called *zones*. These zones are relatively homogeneous areas endowed with a reference health centre, and composed by the aggregation of census enumeration areas defined by IBGE, the Brazilian official census agency. The boundaries of the zones were digitized by the municipal Health Office and they are shown in Figure 1, where some zones are labelled for later reference purposes in the paper. The human cases of HVL are classified according to their occurrence zone and date of first symptoms. In this study, the available information is composed by the number of human disease cases and total population count for each of the 117 zones and each of the periods 1994/1995, 1995/1996 and 1996/1997. A period comprises the second semester of a year, starting in 07/01, and the first semester of the following year, ending in 06/30.

We count the registered cases in each area and in each period. We do not have an estimate of coverage but, HVL being a serious disease requiring careful medical assistance, we believe most of the cases seek the health services. Concerning hospital notification, we also believe it is almost complete, since the medication is delivered by health officials following the notification.

We cannot georeference 2.9 per cent of the cases' data because they are homeless patients. In this situation we do not assign an area to the case and discard it from the analysis.

3. MAPPING DISEASE RATES

The classical procedure to deal with disease event rates is to map the crude rates (CR) or some standardized rate, if age–sex subgroups information is available. The crude rates are

the maximum likelihood estimates of the underlying rates if the counts follow a Poisson distribution. That is, suppose that

$$(X_i|P_i, \xi_i) \sim \text{Poisson}(P_i\xi_i)$$

where X_i is the count of disease cases, P_i is the risk population, and ξ_i is the underlying rate of occurrence in area, i, $i = 1, \ldots, N$. Additionally, we assume that the counts are independent, conditionally on the underlying rates. Then, the maximum likelihood (and uniformly minimum variance unbiased) estimator of ξ_i is given by $\hat{\xi}_i = X_i/P_i$.

However, when the investigation areas (and their populations) are small and the disease under study is rare, the maps made with the CRs can be highly affected by the extreme variability of rates calculated in the less populated areas [11]. In fact, in areas with small population, a small change in the number of cases implies a considerable change in the crude rates. Usually this change is caused by casual random fluctuation unrelated to the underlying risk. This is observed in Figures 2(a)–(c), which present the maps of HVL crude rates in Belo Horizonte. Note that some zones show large changes over time, with some zones standing out at some time points and fading away in the next period. This suggests that the maximum likelihood estimator has large variance because, although possible, it is not likely that these large changes are due to corresponding changes in the underlying risks over the period.

A large literature has been developed to map small area statistics [12–15]. Clayton [16] proposed a hierarchical model which was further developed by Besag *et al.* [17] in a spatial Bayesian setting which has generated much interest. The idea is to impose a plausible spatial relationship structure among the areas and to model the logarithm of the relative risks as an autonormal joint distribution. As a consequence, the information on neighbouring areas is used to improve the estimation in a given area. This procedure avoids the instability of estimates based only on the crude rates of that area, since it uses the information about the events in neighbouring areas which should have similar risks. The definition of neighbourhood is based on the adjacency: areas sharing boundaries are considered neighbours. Other flexible neighbourhood structures have been suggested in the context of disease mapping [18].

Recently, a study [19] was published about the geographical variability of breast cancer and Hodgkin's disease in the Sardinia region, Italy. In the comparison between the maps using the maximum likelihood estimates and those using the Bayesian procedure, the results show that the Bayesian maps are smoother than the crude rates maps, resulting in visually more informative maps with simpler epidemiological interpretation. The Bayesian approach has become the main methodology to map disease rates [11].

4. SPACE–TIME MODELLING OF RATES

There are not many studies addressing the issue of space and time modelling of rates. Most of them deal with epidemics of infectious diseases [20] and they use compartment models concentrating more on the disease dynamics than in its spatial aspects [21]. Much effort has been made to add spatial components to susceptible-infective-removed models, as well as to many linear and non-linear regression formulations [22].

A few studies deal with small area statistics problems for space–time data. Waller *et al.* [23] analysed a data set on mortality from lung cancer for 21 successive years in the 88

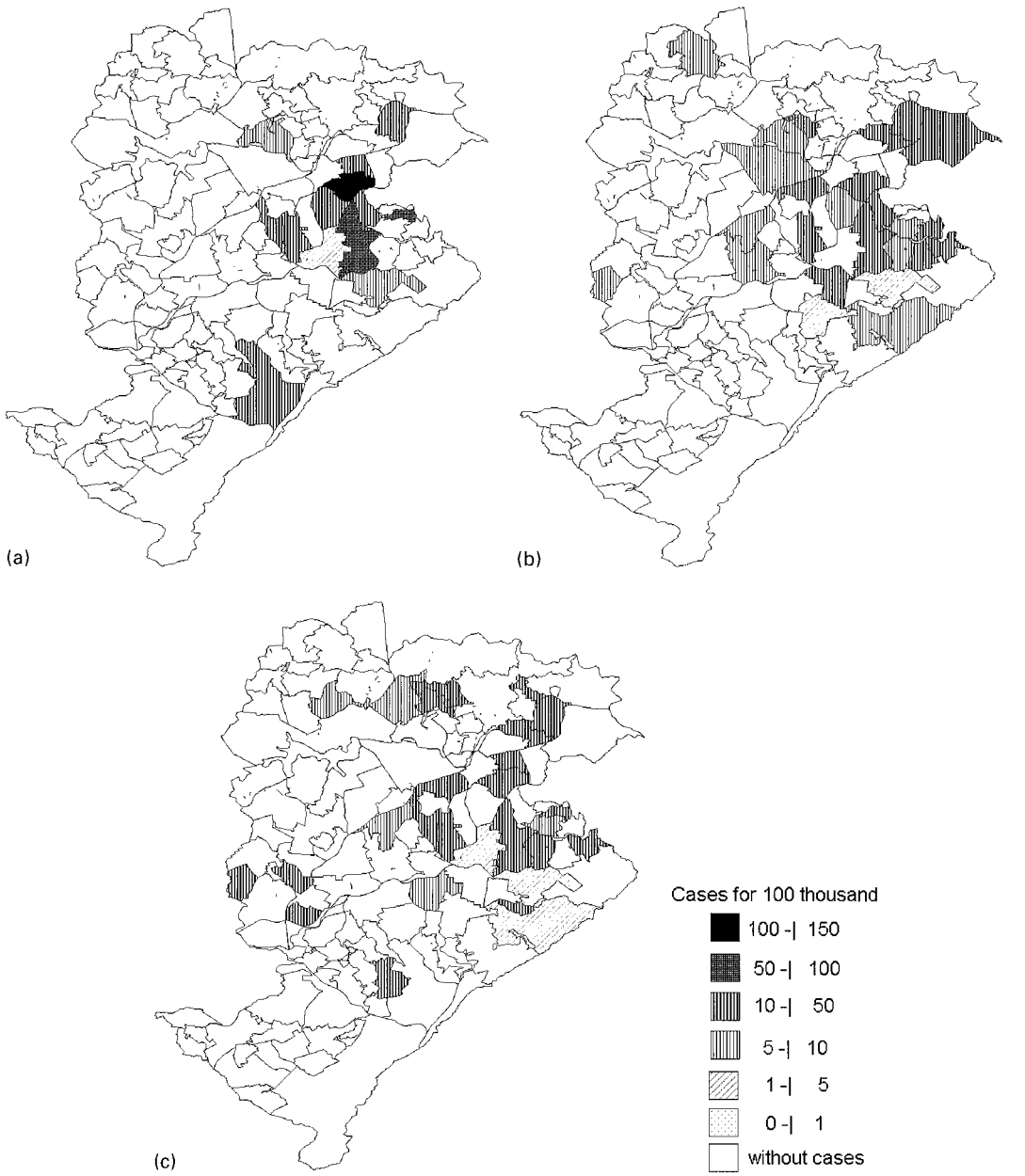

Figure 2. Maps of HLV crude rates (per 100 thousand) in the zones of Belo Horizonte: (*a*) the period 1994/1995; (*b*) the period 1995/1996; (*c*) the period 1996/1997. White areas are those with zero inhabitants.

counties of Ohio. Xia and Carlin [24] and Knorr-Held and Besag [25] reanalysed the same data set, improving the results. We discuss their models in the next section.

Exceptions to the neglect of space–time interaction modelling are the works of Bernadinelli *et al.* [10] and Songini *et al.* [26] on the analysis of the geographical variation of insulin-dependent diabetes mellitus (IDDM) across Sardinia. They model the logarithm of the rates as a linear function of time with area-specific coefficients having a spatial prior which implies neighbouring areas exhibiting similar risk profiles. However, in their application of their model to the IDDM data, they found no evidence of space–time interaction. That is, their final fitted model had each area following a linear trend with area-specific intercept and a constant time-trend coefficient.

In our paper, using a different parameterization, we adopt the model proposed by Bernadinelli *et al.* [10]. In contrast to their results, we find strong evidence of space–time interaction; we use the model to project the rates, and discuss model selection using a cross-validation type of procedure.

To describe our method, denote by $\boldsymbol{X}$ the matrix with element X_{it} representing the number of cases of HVL in area i at time t, $i = 1, \ldots, 117$ and $t = 1, 2, 3$. The time index corresponds to the one year periods from mid-1994 to mid-1997. Similarly, we define the matrix P with element P_{it} representing the population of area i in time t. Finally, we define the matrix E with element ξ_{it} representing the underlying rate of events in area i at time t.

Our model assumes that, given the rates E and the populations P, the counts are independent and

$$(X_i | P_i, \xi_i) \sim \text{Poisson}(P_i \xi_i) \tag{1}$$

This distributional assumption is usually made for vital events and Brillinger [27] derived it through an approach where vital rates are treated as statistics calculated out of Poisson process models generating events on the Lexis diagram.

To track the time evolution of the rates and to make projections, we work with the logarithm of the rates, $\theta_{it} = \ln(\xi_{it})$, rather than the rates directly. This transformed parameter is quite common in statistical studies because it is the natural parameter of the Poisson distribution seen as a member of the exponential family. This brings some theoretical advantages, in terms of the existence of a set of simple sufficient statistics for regression parameters and some simplification of the computational algorithm.

We adopted an area-specific second-degree polynomial trend model for the evolution of the rates over time. This allows acceleration of deceleration of the rates to be captured. Since the projection we have in mind is a short term one, the well known disadvantages of unrealistic long term projections of polynomial models are avoided. Hence

$$\theta_{it} = \ln(\xi_{it}) = \alpha_i + \beta_i(t-1) + \gamma_i(t-1)^2 \tag{2}$$

for $i = 1, \ldots, 117$ and $t = 1, 2, 3$.

In this way, α_i represents the logarithm of the rate in area i in the first year and $\beta_i + \gamma_i(2t-1)$ represents the increment on the logarithm of the area i rate when we pass from year t to $t+1$. We denote by $\boldsymbol{\alpha}, \boldsymbol{\beta}$ and $\boldsymbol{\gamma}$ the vectors with all α_i, β_i and γ_i, respectively.

We want to emphasize that the parameters of the polynomial are specific by area. This implies that each area will have its own estimated regime of increase, stability, or decrease over time. This also implies that we are allowing an interaction between space and time in our procedure since the effect of time will not be the same in different areas.

Each area-specific polynomial has three parameters and we have just three observed counts over time. Hence, there are too many parameters in the likelihood function to make this model useful. To keep this model, it is necessary to impose additional constraints and this comes through the prior distribution in the Bayesian approach, as we describe next.

5. THE BAYESIAN APPROACH

In the Bayesian approach, the prior knowledge about the parameters $\boldsymbol{\alpha}, \boldsymbol{\beta}$ and $\boldsymbol{\gamma}$ can be expressed by a distribution that stimulates configurations with neighbouring areas showing similar values for these parameters. This distribution does not indicate the location of eventual high and low values in a map. Rather, it merely states that the values are dependent and tend to be similar if they are neighbours. A distribution satisfying this requirement has density of the following form:

$$\begin{aligned} f(\alpha_1,\ldots,\alpha_{117}|\tau_\alpha) &\propto \exp\left(-\frac{\tau_\alpha}{2}\sum_{i\sim j}(\alpha_i-\alpha_j)^2\right)\\ f(\beta_1,\ldots,\beta_{117}|\tau_\beta) &\propto \exp\left(-\frac{\tau_\beta}{2}\sum_{i\sim j}(\beta_i-\beta_j)^2\right)\\ f(\gamma_1,\ldots,\gamma_{117}|\tau_\gamma) &\propto \exp\left(-\frac{\tau_\gamma}{2}\sum_{i\sim j}(\gamma_i-\gamma_j)^2\right) \end{aligned} \tag{3}$$

where $i\sim j$ if the areas i and j are neighbours. Note that, by making neighbouring parameter values similar, we decrease the sums in the exponents, increasing the value of the density. Strictly speaking, this is not a proper density since it is invariant by adding a constant to all entries of the vectors $\boldsymbol{\alpha}, \boldsymbol{\beta}$ or $\boldsymbol{\gamma}$. This will not be a problem because the posterior density will be proper even with these improper densities [17]. We also assume that $\boldsymbol{\alpha}, \boldsymbol{\beta}$ and $\boldsymbol{\gamma}$ are independent conditionally on the hyperparameters τ_α, τ_β and τ_γ. Hence, the joint prior distribution is the product of the distributions in (3).

Denote by $\boldsymbol{\alpha}_{-i}$ the vector $\boldsymbol{\alpha}$ with the ith co-ordinate removed. In the same way, define the vectors $\boldsymbol{\beta}_{-i}$ and $\boldsymbol{\gamma}_{-i}$. Considering the values in the vector $\boldsymbol{\alpha}$, denote by $\overline{\alpha_i}$ the arithmetic mean of the n_i neighbouring areas of i. Similarly, define $\overline{\beta_i}$ and $\overline{\gamma_i}$. The distributions in (3) imply conditional *a priori* distributions for α_i, β_i and γ_i expressed, respectively, by

$$\begin{aligned} (\alpha_i|\boldsymbol{\alpha}_{-i}) &\sim N(\overline{\alpha_i}, 1/(\tau_\alpha n_i))\\ (\beta_i|\boldsymbol{\beta}_{-i}) &\sim N(\overline{\beta_i}, 1/(\tau_\beta n_i))\\ (\gamma_i|\boldsymbol{\gamma}_{-i}) &\sim N(\overline{\gamma_i}, 1/(\tau_\gamma n_i)) \end{aligned} \tag{4}$$

This way of expressing the prior knowledge about $\boldsymbol{\alpha}, \boldsymbol{\beta}$ and $\boldsymbol{\gamma}$ takes into account the spatial structure between the areas since the value of one parameter is centred around the mean of their neighbours' values. Distributions defined in (3) are usually called *conditional autoregressive* (CAR).

The degree of similarity among neighbouring area parameters is controlled by the hyperparameters τ_α, τ_β and τ_γ, which are assumed to follow independent gamma distributions. More specifically, $\tau_\alpha \sim \mathbf{\Gamma}(a,b), \tau_\beta \sim \mathbf{\Gamma}(c,d)$, and $\tau_\gamma \sim \mathbf{\Gamma}(e,f)$ where a, b, c, d, e and f are known and, in this paper, set equal to 0.001, which leads to 'only just' proper prior distribution [28]. We comment on possible problems with this prior specification in the Discussion section.

After collecting the data X_{it}, the knowledge about the parameters $\boldsymbol{\alpha}, \boldsymbol{\beta}$ and $\boldsymbol{\gamma}$ is updated through the posterior distribution given by

$$f_{\text{post}}(\boldsymbol{\alpha}, \boldsymbol{\beta}, \boldsymbol{\gamma}, \tau_\alpha, \tau_\beta, \tau_\gamma | X) \propto L(X | \boldsymbol{\alpha}, \boldsymbol{\beta}, \boldsymbol{\gamma}, \tau_\alpha, \tau_\beta, \tau_\gamma) f_{\text{prior}}(\boldsymbol{\alpha}, \boldsymbol{\beta}, \boldsymbol{\gamma}, \tau_\alpha, \tau_\beta, \tau_\gamma) \tag{5}$$

The function $L(X | \boldsymbol{\alpha}, \boldsymbol{\beta}, \boldsymbol{\gamma}, \tau_\alpha, \tau_\beta, \tau_\gamma)$ is the likelihood function of the model in (2) and

$$f_{\text{prior}}(\boldsymbol{\alpha}, \boldsymbol{\beta}, \boldsymbol{\gamma}, \tau_\alpha, \tau_\beta, \tau_\gamma)$$

is the joint prior distribution of the parameters, given by the product of the expressions in (3) times the hyperparameters' densities. Through the posterior distribution of the vectors $\boldsymbol{\alpha}, \boldsymbol{\beta}$ and $\boldsymbol{\gamma}$, we make inferences concerning them, the logarithms of the rates and the rates ξ_{it} themselves.

It is not possible to obtain analytical expressions for this posterior distribution due to the high dimensionality of the problem. One alternative, which has become standard, is to generate samples of the posterior distribution through Markov chain Monte Carlo methods [29]. We used the program BUGS 0.5 to carry out our calculations [28].

We see some important differences between our model and others used in the literature [23–25]. To highlight the differences, we simplify the latter ones eliminating the covariates and age–sex strata. Hence, in our notation, the Xia and Carlin [24] model reduces to $\theta_{it} = \alpha + \beta t + \phi_i^t$ where, given the hyperparameters, the set $\{\phi_1^t, \ldots, \phi_n^t\}$ has a CAR spatial distribution at time t and the sets are independent across time. Therefore, the space–time interaction appears as a spatial structure changing over time. Rather than modelling the time change pattern, they left ϕ_1^t and ϕ_1^{t+1} unrelated in the prior distribution and in the likelihood. The same is true for the Waller *et al.* [23] model which becomes $\theta_{it} = \theta_i^t + \phi_i^t$ where θ_i^t are spatially unstructured parameters and ϕ_i^t follows a CAR distribution within time and both sets, $\{\theta_1^t, \ldots, \theta_n^t\}$ and $\{\phi_1^t, \ldots, \phi_n^t\}$ are independent across time given the hyperparameters. The model proposed by Knorr-Held and Besag [25] reduces to $\theta_{it} = \alpha_t + \theta_i + \phi_i$ where θ_i are spatially unstructured parameters and ϕ_i follows a CAR distribution. Their approach does not model space–time interaction since the effects of time and space are additive.

In contrast, our model allows for space–time interaction in the form of area-specific time trends which varies smoothly in space through the spatial correlation between the trend parameters. We prefer our approach because it mixes the space and time components in a simple way, easy to understand and to explain to epidemiologists. At the same time, it is flexible enough to model short time series in each area, as for our case.

6. MODEL SELECTION

Initially, we model the time evolution of the logarithm of the HVL rates as a second degree polynomial in t, as shown in (2). However, we want to consider the suitability of modelling

this evolution with a first degree polynomial or with a model having time-constant rates. Hence, we propose two additional models to describe the time evolution of the logarithm of the rate. The linear model is given by

$$\theta_{it} = \ln(\xi_{it}) = \alpha_i^* + \beta_i^*(t-1) \tag{6}$$

and the time-constant rates model is given by

$$\theta_{it} = \ln(\xi_{it}) = \alpha_i^{**} \tag{7}$$

$i = 1, \ldots, 117$ and $t = 1, 2, 3$.

In the Bayes approach, one way to compare two models is through the Bayes factor, defined as

$$B = \frac{P(X|M=1)}{P(X|M=2)}$$

where M is the model indicator. It compares the models through the ratio between the marginal density of the observed data under model 1 and under model 2. We can write B as

$$B = \frac{P(M=1|X)P(M=2)}{P(M=2|X)P(M=1)}$$

where $P(M=1)$ and $P(M=2)$ are the prior probabilities of models 1 and 2, respectively, and $P(M=1|X)$ is the posterior probability of model 1. Similarly, we define $P(M=2|X)$.

A procedure to calculate the Bayes factor in complex models has been proposed [30]. The main idea is to obtain $P(M=j|X)$, for $j=1,2$, through the introduction of the indicator M as an additional variable in the problem and, after that, to estimate its posterior probability as the frequency of the $M=1$ event occurs in the MCMC run.

Unfortunately, this procedure has not worked well in practice. The MCMC algorithm did not jump between the models (either quadratic and constant, or quadratic and linear), remaining in only one of them along the chain run. The proponents of this method had the same difficulty [30]. In spite of the adjustments made in $P(M=1)$ and $P(M=2)$ to allow these jumps to occur, as suggested by those authors, the problem persisted in our analysis.

Another alternative to deal with the issue of model selection is to omit some data from the fitting process and then to choose that model which makes the best prediction of those omitted data. For that, we select areas with at least one disease case in the last period. In each one of these areas, one at a time, we omit the observation from the last period and run the MCMC algorithm. The elements of the matrix X with the pseudo missing data are treated as a parameter, and a sample is obtained from the posterior distribution of this new parameter. The justification to not select areas with zero cases in the last year is that most of the predictions in these cases would be equal to the observed value, zero, and hence less informative about the different performances of the models.

Out of a good model, we expect to have the posterior distribution of the parameter centred in the observed value of the corresponding data. Hence, we define a measure p as

$$p = \min\{P(X_{it} < v_{it}|X_{-it}),\ P(X_{it} > v_{it}|X_{-it})\}$$

where v_{it} is the true value of X_{it} and X_{-it} is the matrix of cases counts with the value v_{it} missing. The measure p varies between 0 and 0.5, and the larger p, the better the model.

Table I. Descriptive statistics for the p measure of the proposed models for the logarithm of human visceral Leishmaniasis in 11 areas of Belo Horizonte.

Model	Median	Mean	1st and 3rd quartiles	Coefficient of variation
Constant	0.047	0.077	0.033–0.101	0.919
Linear	0.120	0.174	0.086–0.263	0.722
Quadratic	0.105	0.136	0.065–0.191	0.670

Using the posterior distribution sample, we can calculate an estimate of p, with the above probabilities evaluated through the empirical frequencies in the MCMC run.

7. RESULTS

The three proposed models to describe the time evolution of HVL rates (constant, linear and quadratic) are compared through the procedure described in the Section 6, evaluating the measure p in 11 areas independently chosen among those areas where there was at least one case recorded during the study period. The measure p is calculated using the estimated *a posteriori* mean ξ_{-i3} distribution of ξ_{i3} obtained with the information of the last period in area i missing. This is done in the following way:

$$p = \min\{P(W_{i3} < v_{i3}|X_{-i3}),\ P(W_{i3} > v_{i3}|X_{-i3})\}$$

where W_{i3} is a random variable with Poisson distribution with mean $\xi_{-i3}P_{i3}$. The results are presented in Table I using 20 000 iterations collected every 10th element with 10 000 iterations as burn-in.

Analysing Table I, we note that we should choose one between the linear and the quadratic models. To simulate from the posterior distribution of the underlying rates θ_{it}, we substitute in (6) and in (2) the MCMC generated samples of α^*, β^* and of $\boldsymbol{\alpha}, \boldsymbol{\beta}$ and $\boldsymbol{\gamma}$, respectively. In this way we estimate the HVL rates posterior distribution in each zone for the three years studied. We can also find the posterior distribution of the predicted θ_{i4} value. We summarize the posterior distributions by the posterior mean $E(\theta_{it}|X)$.

Figure 3(b) shows the plot of the crude rates for the 40 zones with at least one case in the three periods. Figure 3(a) shows the plot of $E(\theta_{it}|X)$, $t = 1, 2, 3, 4$, using the linear model for these same 40 areas and Figure 3(c) uses the estimates based on the quadratic model. We note that, using crude rates, we can neither visualize well the time evolution of the rates nor notice the areas with higher risks of occurrence of the disease, except for areas 55, 43, 36 and 64, which stand out very clearly. The estimates of the other zones suffer from the high instability of the crude rates, varying abruptly from year to year. Analysing the graphs with the rates estimated by the linear and quadratic models, with the projections for the following period, we can notice that, in both models, some areas stand out due to their risk profile. These areas are labelled by their zone code in the plots. The effect of the Bayesian estimates is to make clearer the time evolution of the rates, pointing out areas which can be of great concern.

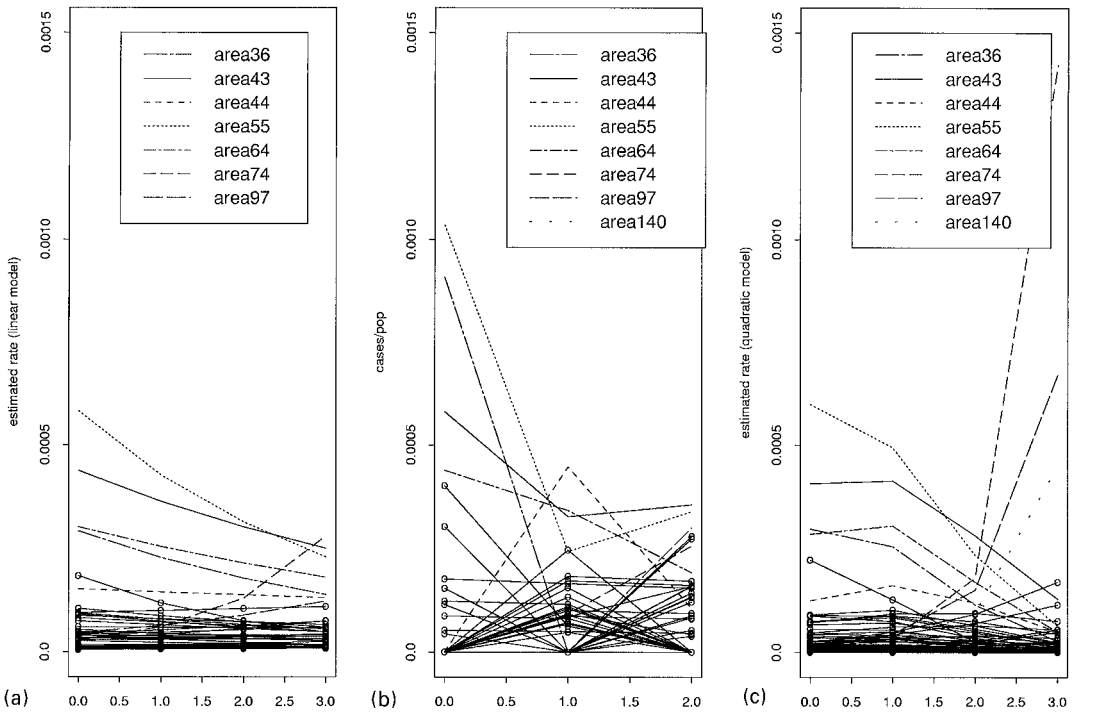

Figure 3. Time plot for the rates of 40 zones with at least one case in the three periods. (*a*) Plot of the Bayesian estimates of the rates using the linear model for the logarithm of the rates. (*b*) Plot of the crude rates. (*c*) Plot of the estimates based on the quadratic model. Time 0 is the period 1994/1995, time 1 is the period 1995/1996, and time 2 is the period 1996/1997. In (*a*) and (*c*), time 3 corresponds to the projected 1997/1998 period.

Comparing Figures 3(*a*) and 3(*c*), we see that, in general, the projections given by the quadratic model are more pessimistic than those from the linear model, except by area 74, which appears as the area with the fastest increase. The quadratic model highlights the same areas as the linear model, and some additional ones, such as areas 97 and 74. These latter areas are not close to those where the disease was introduced in Belo Horizonte, especially area 74, located in the opposite side of the municipality. This fast increase can characterize a new disease focus.

The objective of the projections is to provide information to allocate public efforts in the areas affected by the disease and to evaluate the results of these efforts. Therefore, extremely high precision to the estimates and projections is not crucial if the decisions are the same whatever the adopted model, as it is the case here. Both models show in a clear way which areas are worrisome and how the disease is developing in time. Any model leads to approximately the same ranking and the same relative priorities among the areas.

Figures 4(*a*) to (*c*) and Figure 5 show the maps using the rates estimated by the quadratic model for the three periods considered and the projection for the next period. Comparing with the maps of Figure 2, we note that the Bayesian maps are smoother and show the spread of the disease from Northeast and East in direction to the Northeastern and Northern regions.

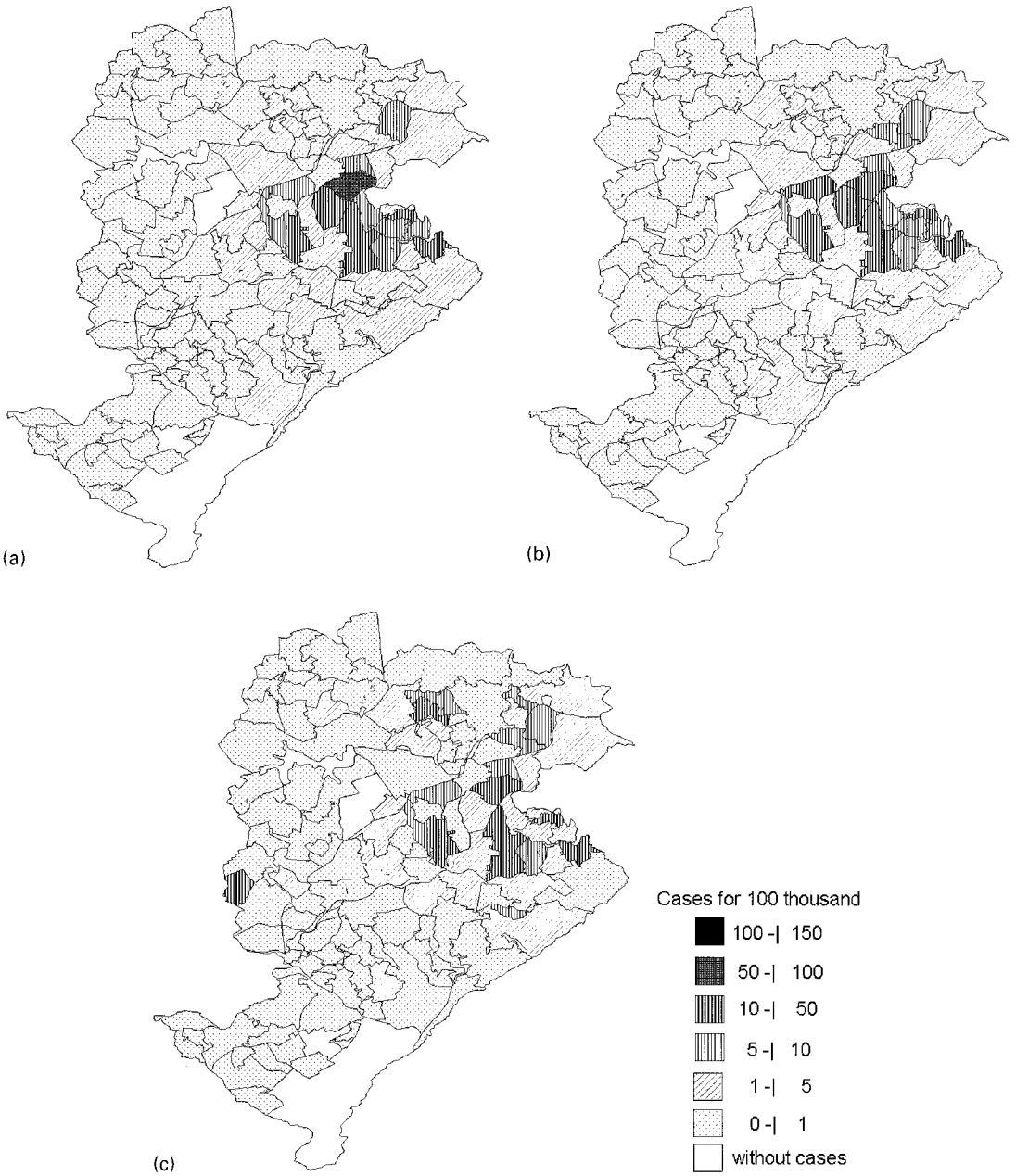

Figure 4. Maps of MCMC posterior mean estimates of HLV rates (per 100 thousand) in the zones of Belo Horizonte (*a*) the period 1994/1995; (*b*) the period 1995/1996; (*c*) the period 1996/1997. White areas are those with zero inhabitants.

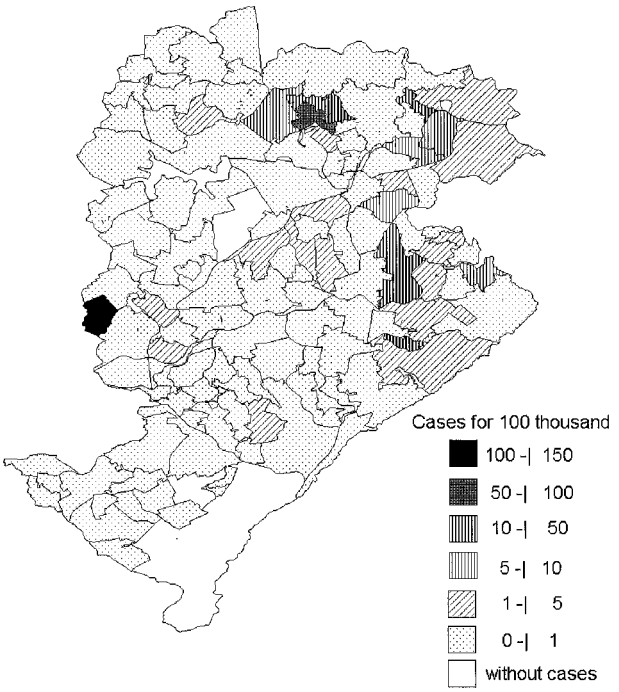

Figure 5. Map for the Bayesian projections of human visceral Leishmaniasis rates (per 100 thousand) in the zones of Belo Horizonte for the period 1997/1998. White areas are those with zero inhabitants.

Figure 6 shows the sample trace plots and kernel posterior density estimate of the hyperparameters $\tau_\alpha, \tau_\beta, \tau_\gamma$. Observe that the plots have different scales. The τ_β and τ_γ plots provide further evidence for the space–time interaction. In fact their 95 per cent credibility intervals are (0.78, 17.90) and (2.47, 35.40) and, since the τ_γ interval is much larger than that of τ_β, there is evidence that the space–time interaction occurs primarily in the linear trend rather than the quadratic time trend. The reason is the typically smaller precision and, therefore, the larger variance of the linear trend parameters across areas as compared with the quadratic term. To clarify this statement, consider what are the typical realizations from the posterior joint distribution of β and of γ. Because there is appreciable posterior probability that $\tau_\beta < \tau_\gamma$, the map of a typical realization of γ will be smoother than that of β. Hence, the changes on the time profile from area to area are due more substantially to differences in β rather than to differences in γ.

We tested convergence of the generated MCMC runs checking the hyperparameters and of some area-specific parameters posterior distribution. Based on Geweke's [31] convergence diagnostic test, we calculated the Z-score statistic for the three hyperparameters $\tau_\alpha, \tau_\beta, \tau_\gamma$, finding 1.18, -1.27 and 0.995, respectively. Since these statistics should have standard normal distribution under the null hypothesis of stationarity, we find no evidence against convergence of the MCMC runs. In the Heidelberg and Welch [32] test, we had the Cramer–von Mises statistic with values 0.279, 0.320 and 0.280 for hyperparameters $\tau_\alpha, \tau_\beta, \tau_\gamma$, respectively, again providing no evidence of non-stationarity.

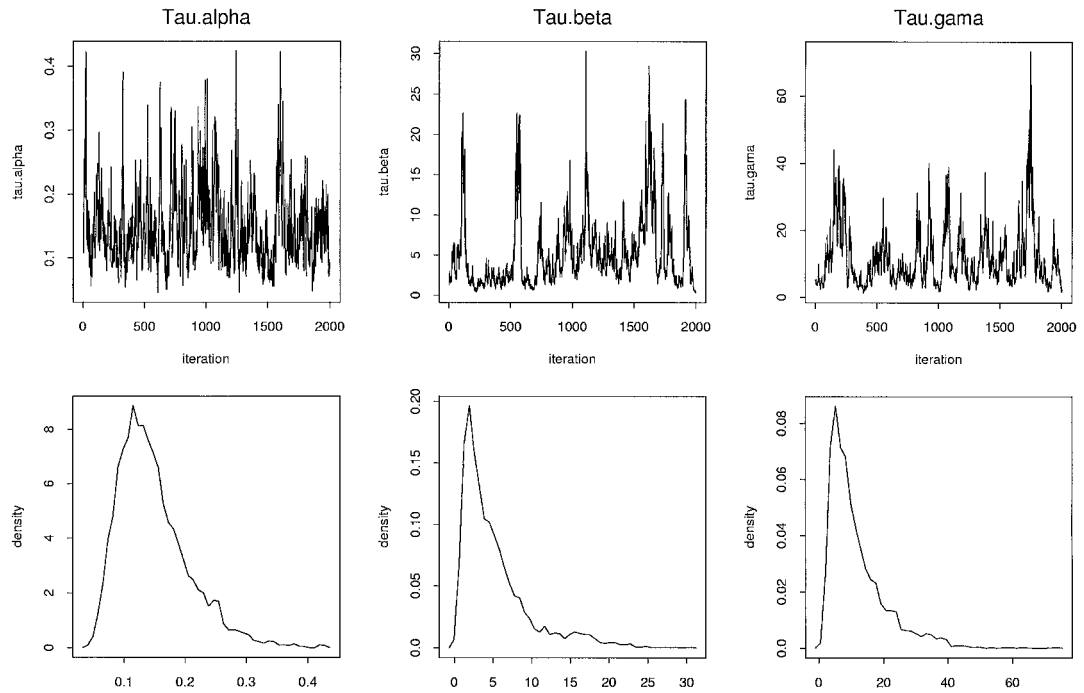

Figure 6. Sample trace plots and kernel posterior density estimate of the hyperparameters $\tau_\alpha, \tau_\beta, \tau_\gamma$ for the quadratic model.

We also calculated the Gelman and Rubin [33] R convergence statistic based on four quite different initial values for all parameters. Monitoring the hyperparameters $\tau_\alpha, \tau_\beta, \tau_\gamma$ in batches of 50 updates, we found the median R to be equal to 1.01, 1.09, 0.97 and the 97.5 per cent percentile to be 1.16, 1.12, 1.08, respectively, and therefore, the chains converged according to this criterion. We also calculated the above tests using the α_i, β_i and γ_i chains from five randomly selected areas, again finding no evidence of non-stationarity.

8. DISCUSSION

In this article we make inferences for the rates of human visceral Leishmaniasis in Belo Horizonte from 1994 to 1997, with projection to June 1998. We use a methodology based on the Bayesian approach fitted with the help of MCMC methods. This procedure filters out a large part of the variability due to the random fluctuation in the crude rates, a fluctuation which is not associated with the underlying risks. Additionally, the methodology incorporates in an explicit way the discrete nature of the counts and, implicitly, the spatial correlation between underlying rates of neighbouring areas. The latter is obtained through the spatial structure of the parameters' trend on time in each area. We also propose a method to choose different models for the risk evolution.

The gamma hyperpriors we adopted are proper and it is suggested in the BUGS software and elsewhere that they are nearly non-informative. However, as one referee pointed out, there is room for concern in adopting this kind of hyperprior since Natarajan and McCulloch [31] have recently shown that its use in some examples can lead to inaccurate posterior estimates which are not be identified from the MCMC runs. For a class of spatial models, related but not containing those used in this paper, sufficient and necessary conditions for a proper posterior has been given [32, 33]. It is not obvious how to generalize these results for our specific model and the authors are currently working on this issue.

The Bayesian results provide a clear picture of the complex disease development. The information on Figures 3 and 4 suggests that the health officials' actions avoided the increase of the epidemic curve since most of the rates decline in time. The decline was steeper on areas with initially high prevalence. However, the public health actions were not sufficient to prevent the geographical diffusion of the disease and the appearance of new foci of the disease. The declining rates coupled with diffusion can be seen in the sequence of Bayesian maps which show a final picture of few and relatively small disease foci scattered in the region. The Eastern focus is present since the beginning of the period, the Northern focus seems to appear by diffusion from the initial one and the Western focus is probably a genuinely new focus.

We believe that the control measures adopted in Belo Horizonte by the municipal Zoonosis Service contributed to reduce the human incidence. In 1994, 3402 houses were sprayed with pyrethroid insecticides and this number increased to 46 980 in 1995. Next year, there were 46 258 sprayed houses, and in 1997 this number reduced to 22 640 houses. Additionally, in part of the city during 1996, the Zoonosis Service used spatial spraying with organophosphate insecticide spray for *Aedes aegypti* because of one dengue epidemic and this action is likely to have reduced the Leishmaniasis visceral vector population.

The geographical diffusion of the disease can be explained by two factors. The first one is the dog search procedure. The service considers only those dogs living in households and we suspect there is a relatively large number of stray dogs in the city. Positive untested dogs wandering in the city can spread the disease relatively fast. Another possible factor is the intense action of the officials during the years of this study in areas with high canine prevalence. This could favour the increase of the disease in the other areas with low prevalence and therefore lower control.

The adoption of the methodology of spatial Bayesian models can contribute to redirect the control programme, with larger priority to areas where the Bayesian projections indicate larger increase. Let us point out, however, that the follow-up period is short and the secondary nature of our data is subject to the typical biases of this type of information, such as, for example, the underreporting of cases.

The results show that, as opposed to the crude rates, the Bayesian estimates are much simpler and clearer to use to orient public policy decisions in this health problem. They provide a better scenario of areas with increasing or stable rates, a classification of areas in high and low risk zones, and they point out the presence of a new disease focus which needs close scrutiny.

ACKNOWLEDGEMENTS

This research was supported by FAPEMIG, CNPq and FUNDEP. The authors thank Peter Diggle, Trevor Bailey, David Spielgelhater and Michel Spira for suggestions and comments on this work. We also want to thank the referees for comments and suggestions that substantially improved this paper.

REFERENCES

1. Chagas E, Cunha AM, Ferreira LC. Leishmaniose Visceral Americana. *Memórias do Instituto Oswaldo Cruz* 1938; **33**:89–229.
2. Deane LM, Deane M. Observações preliminares sobre a importância comparativa do homem, do cão e da raposa (*Lyealopex vetulus*) como reservatório da *Leishmania donovani* em área endêmica de calazar no Ceará. *O Hospital* 1955; **48**:61–76.
3. Alencar JE. Profilaxia do Calazar no Ceará, Brasil. *Revista do Instituto de Medicina Tropical de São Paulo* 1961; **3**:175–180.
4. Secretaria Municipal de Saúde de São Luís, MA. Programa Emergencial do Controle da Leishmaniose Visceral em São Luís/MA.SMSSL, 1995.
5. Jerônimo SMB, Oliveira RM, Mackay S. Urban outbreak of visceral leishmaniasis in Natal, Brazil. *Transactions of the Royal Society of Tropical Medicine and Hygiene* 1994; **88**:386–388.
6. Monteiro PS, Lacerda MM, Arias JR. Controle da Leishmaniose Visceral no Brasil. *Revista da Sociedade Brasileira de Medicina Tropical* 1994; **27**(suplemento III):67–72.
7. Costa CHN, Pereira HF, Araújo MV. Epidemia de Leishmaniose Visceral no Estado do Piauí, Brasil, 1980–1986. *Revista de Saúde Pública* 1990; **24**:361–372.
8. Genaro O, Costa CA, Willins P, Silva JE, Rocha NM, Lima SL. Ocorrência de calazar em área urbana da grande Belo Horizonte, MG. *Revista de Sociedade Brasileira de Medicina Tropical* 1990; **23**:67–72.
9. Cosenza GW. Leishmaniose Visceral em Belo Horizonte. *Boletim Epidemiológica SUS-MG* 1995; **4**:2–12.
10. Bernadinelli L, Clayton D, Pascutto C, Montomoli C, Ghislandi M, Songini M. Bayesian analysis of space–time variation in disease risk. *Statistics in Medicine* 1995; **14**:2433–2443.
11. Mollié A. Bayesian mapping of disease. In *Markov Chain Monte Carlo in Practice.* Gilks WR, Richardson S, Spielgelhalter DJ (eds). Chapman & Hall: London, 1996: 359–379.
12. Tsutakawa RK. Mixed model for analysing geographic variation in mortality rates. *Journal of the American Statistical Association* 1988; **83**:37–42.
13. Tsutakawa RK, Shoop GL, Marienfeld CJ. Empirical Bayes estimation of cancer mortality rates. *Statistics in Medicine* 1985; **4**:201–212.
14. Manton KG, Woodbury MA, Stallard E, Riggan WB, Creason JP, Pellom AC. Empirical Bayes procedures for stabilizing maps of U.S. cancer mortality rates. *Journal of the American Statistical Association* 1989; **84**: 637–650.
15. Clayton DE, Kaldor J. Empirical Bayes estimates of age-standardized relative risks for use in disease mapping. *Biometrics* 1987; **43**:671–681.
16. Clayton D. Hierarchical Bayesian models in descriptive epidemiology. *Proceedings of the XIVth International Biometrics Conference* 1989; 201–213.
17. Besag J, York J, Mollié A. Bayesian image restoration, with applications in spatial statistics (with discussion). *Annals of the Institute of Statistical Mathematics* 1991; **43**:1–59.
18. Conlon E, Waller LA. Flexible neighborhood structures in hierarchical models for disease mapping. Research report 98-018, University of Minnesota, Department of Biostatistics.
19. Bernadinelli L, Montomoli C. Empirical Bayes versus fully Bayesian analysis of geographical variation in disease risk. *Statistics in Medicine* 1992; **11**:983–1107.
20. Cliff AD, Haggett P. *Atlas of Disease Distributions.* Blackwell: Oxford, 1988.
21. Grenfell BT, Kleczkowski A, Gilligan CA, Bolker BM. Spatial heterogeneity, nonlinear dynamics and chaos in infectious diseases. *Statistical Methods in Medical Research* 1995; **4**:160–183.
22. Cliff AD, Haggett P. Statistical modelling of measles and influenza outbreaks. *Statistical Methods in Medical Research* 1993; **2**:43–73.
23. Waller LA, Carlin BP, Xia H, Gelfand AE. Hierarchical spatio-temporal mapping of disease rates. *Journal of the American Statistical Association* 1997; **92**:607–617.
24. Xia H, Carlin B. Spatio-temporal models with errors in covariates mapping Ohio lung cancer mortality. *Statistics in Medicine* 1998; **17**:2025–2043.
25. Knorr-Held L, Besag J. Modelling risk from a disease in time and space. *Statistics in Medicine* 1998; **17**: 2045–2060.
26. Songini M, Bernadinelli L, Clayton D, Montomoli C, Pascutto C, Ghislandi M, Fadda D, Bottazzo GF, and the Sardinian IDDM Study Groups. The Sardinian IDDM Study: 1. Epidemiology and geographical distribution of IDDM in Sardinia during 1989 to 1994. *Diabetologia* 1998; **41**:221–227.
27. Brillinger D. The natural variability of vital rates and associated statistics. *Biometrics* 1986; **42**:693–734.
28. Spiegelhalter DJ, Thomas A, Best NG, Gilks W. *BUGS*: *Bayesian inference using Gibbs sampling*, *Version 0.50*. Medical Research Council Biostatistics Unit: Cambridge, 1996.
29. Gelman A, Carlin JB, Stern HS, Rubin DB. *Bayesian Data Analysis.* Chapman and Hall: London, 1995.

30. Carlin BP, Chib S. Bayesian model choice via Markov chain Monte Carlo methods. *Journal of the Royal Statistical Society*, *Series B* 1995; **57**(3):473–484.
31. Geweke J. Evaluating the accuracy of sampling-based approaches to calculating posterior moments. In *Bayesian Statistics 4*. Bernardo JM, Berger JO, Dawid AP, Smith AFM (eds). Clarendon Press: Oxford, U.K., 1992.
32. Heidelberger P, Welch L. Simulation run length control in the presence of an initial transient. *Operations Research* 1983; **31**:1109–1144.
33. Gelman A, Rubin DB. Inference from iterative simulation using multiple sequences. *Statistical Science* 1992; **7**:493–497.
34. Natarajan R, McCulloch CE. Gibbs sampling with diffuse proper priors: a valid approach to data-driven inference? *Journal of Computational and Graphical Statistics* 1999; **7**:267–277.
35. Ghosh M, Natarajan K, Waller LA, Kim D. Hierarchical Bayes GLM's for the analysis of spatial data: an application to disease mapping. *Journal of Statistical Planning and Inference* 1999; **75**:305–318.
36. Sun D, Tsutakawa RK, Speckman PL. Bayesian inference for CAR(1) models with noninformative priors. *Biometrika* 1999; **86**:314–350.

