Tuberculosis in intra-urban settings: a Bayesian approach

Wayner V. Souza[1], Marilia S. Carvalho[2], Maria de Fátima P. M. Albuquerque[3], Christovam C. Barcellos[4] and Ricardo A. A. Ximenes[5]

1 Aggeu Magalhães Research Centre, Oswaldo Cruz Foundation, Recife, Brazil
2 National School of Public Health, Oswaldo Cruz Foundation, Rio de Janeiro, Brazil
3 Aggeu Magalhães Research Centre, Oswaldo Cruz Foundation/Federal University of Pernambuco, Recife, Brazil
4 Scientific Information Centre, Oswaldo Cruz Foundation, Rio de Janeiro, Brazil
5 Federal University of Pernambuco/State University of Pernambuco, Recife, Brazil

Summary

OBJECTIVE To model the effect of socio-economic deprivation and a few transmission-related indicators of the tuberculosis (TB) incidence at small area level, to discuss the potential of each indicator in targeting places for developing preventive action.

METHODS Ecological spatial study of TB incidence in Olinda, a city in the north-east of Brazil, during the period 1996–2000. Three socio-economic indicators (mean number of inhabitants per household; percentage of heads of household with <1 year's formal education; percentage of heads of households with monthly income lower than the minimum wage) and two transmission-related indicators (number of cases of retreatment; number of households with more than one case during the period under study), all calculated per census tract, were used. We adopted four different full hierarchical Bayesian models to estimate the relative risk of the occurrence of TB via Markov chain Monte Carlo.

RESULTS The best specified model includes all the selected covariates and the spatially structured random effect. The gain in goodness-of-fit statistic when the spatial structure was included confirms the clustered spatial pattern of disease and poverty. In this model, the covariates within the non-zero credibility interval were the number of persons per house, the number of cases of retreatment and the number of households with more than one case (all with relative risk ≥1.8) in each census tract.

CONCLUSIONS The possibility to estimate in the same framework both the contribution of covariates at ecological level and the spatial pattern should be encouraged in epidemiology, and may help with establishing Epidemiological Surveillance Systems on a territorial basis, that allows rational planning of interventions and improvement of the Control Programme effectiveness.

keywords Tuberculosis, Bayesian methods, Spatial Analysis, Deprivation, Epidemiological surveillance

Introduction

The relationship between tuberculosis (TB) and socio-economic status is well known (Souza et al. 2000; Waaler 2002). Besides, recent studies succeeded in demonstrating, at area level, association between poverty and TB incidence rates (Krieger et al. 2003). The objective of this paper is to model the relationship between TB incidence at small area level (census tracts), poverty and measures of the control programme performance, discussing the capability of various indicators in guiding preventive action.

Deprivation, household clustering and lack of health service access are well-known factors associated with TB incidence and should be incorporated in health policies planning. An effective programme for TB control should include these characteristics not only as individual risk factors but as group determinants, composing the basis for

a territorial approach to a surveillance system (Barr et al. 2001).

Strong evidences for a relationship between TB and poverty are already available, expressed by higher TB incidence rates in crowed urban areas and amongst low income and illiterate populations (Waaler 2002). In addition, health services access is impaired in the same populations, both due to lack of adequate services in poor areas and because poorer and illiterate people are less aware of their own health status.

In usual clinical practice, and most epidemiological studies, socio-economic determinants of disease are included in terms of an individual person's risk factors (Kawachi & Berkman 2003). In ecological studies, on the other hand, the focus is on the community as an entity in itself, an entity more complex than the sum of the individual persons who make it up (Berkman & Kawachi

2000). Although individual level factors such as behaviours and lifestyles, common to individuals across areas and subject to public health interventions, may also play a role in TB incidence, the question, in ecological analysis, is not about the causes of disease cases, but the causes of disease incidence (Rose 2001).

Spatial analysis has recently been included as a powerful public health tool, for its capability of visualizing disease distribution, even with sparse data (Best *et al.* 2005), and mapping risk factors at population level (Bailey 2001; Bailey *et al.* 2005). A few recent publications have emphasized the role of Geographical Information Systems technology in Public Health research and practice, its use being limited by the quality of address information, particularly those available in routine information systems. In spite of that, this is the data that should be used when dealing with services and programme evaluation, using the smallest possible area, especially in places where social contrasts are high, as in most developing countries (Camara & Monteiro 2001; Cromley *et al.* 2002; Ricketts 2003).

Nevertheless, spatial analyses based on small areas, remains somewhat limited to statistical literature in spite of recent methodological advances. This may be due to the inherent complexity of solving two problems at the same time: smoothing out the statistical instability of small numbers often present in area data with enough resolution to discriminate differences in socio-economic status; and estimating the effects of covariates in a generalized linear model framework. Moreover, spatial dependence among neighbouring areas increases to the same extent that spatial resolution increases: smaller areas are more similar to each other, and this dependence declines with distance (Bailey 2001; Elliott *et al.* 2001). Bayesian inference has brought important contribution to this field of research which, together with computation power and friendlier software, can be used to public health purposes (Lawson 2001). To understand the true underlying geographical distribution of disease incidence, together with better effects estimation of indicators at ecological level, is helpful when targeting public health action (Militino *et al.* 2001).

Methods

This is an ecological spatial study of TB incidence in Olinda, a city in the north-east of Brazil. With a population of 367 902 spread across 40.83 km^2 (9000/km^2), Olinda is one of the most densely populated cities in Brazil. The spatial unit was the census tracts (the geographic unit used to organize the field work in Brazilian demographic census, composed of approximately 300 households in urban areas), 299 in total. Socio-economic data and the digital map refer to the demographic census of 2000 (Instituto Brasileiro de Geografia e Estatística 2001).

The source of data for TB cases was the Disease Surveillance Warning System, a national system which provides information on all diseases of compulsory notification, including TB. Therefore, the diagnosis of TB was that done by the TB control programme, assuming that the identification of cases by the control programme is good enough to be used as the mechanism of ascertainment of cases in the study, and considering that notification of TB cases is compulsory in Brazil. Diagnosis is mainly based on sputum smear examination of patients presenting with cough lasting more than 3 weeks. Because of the high proportion of smear negative TB cases (sensitivity of only 50–60%) (Aber *et al.* 1980), diagnosis may also be done based on the clinical and epidemilogical evidence. The data covered the period 1996–2000. We used the following indicators, all calculated per census tract: x_1: mean number of inhabitants per household; x_2: percentage of heads of household with <1 year's formal education; x_3: percentage of heads of households with monthly income lower than the minimum wage; x_4: number of cases of retreatment; x_5: number of households with more than one case during the period under study. We used the logarithms of all covariates to normalize the distributions.

Tuberculosis cases are registered in the routine information system used by the official programme of TB control, from 1996 to 2000. Geocoding used the address registered in the dataset, a map and the description of census tracts. In approximately 40% of the cases, the area was visited to localize the cases, as the quality of addresses is characteristically a problem in deprived areas.

We used a generalized linear mixed model (GLMM) to smooth out the variability in observed disease rates and to estimate the association between TB incidence rate, averaged over the observed period and chosen covariates. The expected number of cases in each census tract (ε_i) can be estimated as the municipality overall mean rate (*mean rate*) multiplied by the population in each tract ($\varepsilon_i =$ mean rate $\times$ pop$_i$), and the relative risk – λ_i – of each area is the observed number of cases divided by the expected. In the Bayesian framework, however, the observed number of cases y = (y_1,...,y_n) in the *n* areas is one realization of Poisson random variables, (y_1,...,y_n), with means, $\mu = (\mu_1,...,\mu_n)$. This value can be expressed as number of expected cases ε_i multiplied by the relative risk of each area: $\mu_i = \varepsilon_i \times \lambda_i$, or, in logarithm form, log $\mu_i = \log \varepsilon_i + \log \lambda_i$. The relative risk λ_i is a function of the *k* covariates x_k that 'explain' differences in disease incidence:

$$\lambda_i = \exp(\beta_0 + \beta_{1i}x_{1i} + \beta_{2i}x_{2i} + \cdots + \beta_{ki}x_{ki} + \theta_i + v_i)$$

For β_{ij} we assumed non-informative Gaussian priors with zero mean and precision equal to 1×10^{-5}. In that context, v_i is a non-spatially structured random effect, typically assumed to be independent Gaussian, with zero mean and variance σ_v^2, usually included in the models to account for extra-poisson variation, because of non-measured important covariates. The spatially structured random effects – $\theta = (\theta_1, \ldots, \theta_n)$ – accounts for the spatial dependence, with prior distribution taken as a conditional intrinsic Gaussian autoregressive model, where the mean value for θ_i is an weighted average of the neighbouring random effects and the variance, σ_{θ^2}, controls the strength of this local spatial dependence.

$$p(\theta_i / \theta_{j \neq i}) \sim N\left(\frac{\sum_{j \neq i} w_{ij}\theta_j}{\sum_{j \neq i} w_{ij}}, \frac{\sigma^2}{\sum_{j \neq i} w_{ij}} \right).$$

As most studies using areas, we defined neighbourhood as adjacent census tracts with simple binary adjacency weights, i.e. $w_{ij} = 1$ if area i shares a common boundary with area j and $w_{ij} = 0$ otherwise (Bernardinelli *et al.* 1995; Best *et al.* 2005).

In the first model we fitted just socio-economic indicators, considered to be the most distal determinants: inhabitants per household, income and schooling (distal model). The second model included covariates related to the transmission of the disease: retreatment and more then one case per household (proximal and distal model). In the third one, both random effects (v_i and θ_i) were included (proximal and distal random effects model), and the forth and last model only the random effects were included (null spatial random effects model).

The parameters were estimated via Markov chain Monte Carlo, using the public domain software package Win-BUGS (Cambridge, UK) (Gilks *et al.* 1996; Spiegelhalter *et al.* 1997). Five-hundred thousand simulations were generated, with a burn-in of 150 000 and selecting just one in each 50 samples, to avoid autocorrelation. The mean and 90% credibility intervals were calculated with 7000 posterior samples. Convergence analysis was made in S-Plus (MathSoft, Seattle, WA, USA), using Bayesian output analysis routine, now gives a comparison of the full model. Models were compared using the Deviance Information Criterion (DIC), that combines a measure of fit and a measure of model complexity based on the effective number of parameters. Smaller values of DIC indicate a more appropriate model (Spiegelhalter *et al.* 2002). The WinBugs code used are presented in the appendix.

Results

Olinda had 1984 registered cases of TB during the study period, with a mean incidence of more than one case per one thousand residents. We were able to localize 1678 in the census tract polygons (84.6%), losses because of the poor quality of the addresses.

In Table 1 the measures of association and respective credibility intervals are shown, for all covariates in all models. In distal model, the most important variable was the number of inhabitants per household, indicating that an increase of one more person in the house more than

Table I Estimated risk factors effects associated with tuberculosis incidence by model fitted, Olinda, 1996–2000

Covariate	Distal model (1) Median	Distal model (1) 5–95%	Proximal and distal model (2) Median	Proximal and distal model (2) 5–95%	Proximal and distal random effects model (3) Median	Proximal and distal random effects model (3) 5–95%	Null spatial random effects model (4) Median	Null spatial random effects model (4) 5–95%
Intercept	0.10	0.06–0.17	0.18	0.11–0.30	0.18	0.08–0.38	0.83	0.77–0.88
Persons/house	**2.46**	**1.65–3.63**	**1.67**	**1.12–2.46**	**1.94**	**1.12–3.50**	–	–
No schooling	0.97	0.86–1.09	0.99	0.88–1.12	1.04	0.86–1.24	–	–
Low income	**1.49**	**1.29–1.71**	**1.32**	**1.15–1.52**	1.17	0.95–1.44	–	–
Retreatment	–	–	**1.74**	**1.55–1.95**	**1.80**	**1.47–2.20**	–	–
Another case	–	–	**2.04**	**1.84–2.25**	**1.89**	**1.57–2.27**	–	–
Non-structured random effect variance	–		–		0.029		0.041	
Spatial random effect variance	–		–		0.049		0.153	
DIC	1725.3		1538.0		1409.3		1434.6	
Sample	7000		7000		7000		7000	

DIC, Deviance Information Criterion.
Values in bold refer to the 90% credibility intervals which do not include the unit.

doubled the risk of TB. Schooling, one of the most widely used socio-economic indicator in Brazil, was not important in any model. It was possibly due either to indirect effect of schooling in TB occurrence, through its effect on income (Singh-Manoux *et al.* 2002), or to the cut-off point adopted to define low education. In the models without random effects (models 1 and 2), the effect of income was important: each additional 1% of heads of the household with low income in each census tract increased the risk of TB in 30% (model 2). The largest effect, however, was the number of persons per house. Although an indicator of poverty, it is also important in the household transmission of TB, and is possibly the reason it was the unique socio-economic indicator (the credibility interval does not include the unit) in all models. The inclusion of number of cases of retreatment and the number of households with more than one case in each census tract decreased the effect of the previous covariates.

Adjusting the GLMM (proximal and distal spatial random effects model) affected substantially only the effect of income and, as expected, decreased the effect of all covariates. This result can be understood not only as a consequence of the colinearity between the mean number of persons per house and low income (Pearson's $r = 0.55$) but additionally because of the fact that poverty is spatially clustered. The most important change, however, was in the DIC, considerably smaller than the previous ones. Null spatial random effects model (4) was fitted to disclose the underlying spatial process, as no part of the incidence is explained by any covariate. The DIC of model 4 was a somewhat larger than the model 3.

The comparison of estimated and observed relative risk values is a valuable tool in assessing the fit of the model. Figure 1 shows the improvement of fit in models 3 and 4 as compared with models 1 and 2 because of the inclusion of random effects.

As overdispersion is a common problem with Poisson modelling, alternatively we verified the results of a model with only the non-structured random effect and the five covariates included in model 3. The risk estimates obtained were: intercept (0.16); persons/house (1.76); no schooling (0.99); low income (1.32); retreatment (1.86); another case (2.11). The DIC was 1426.5 and Variance (v) = 0.078, suggesting the need to include the spatially structured random effect to improve the results.

Figure 2 shows the location of cases of retreatment and/or families with more than one case pointed over the density of person per house. These variables presented the largest effect (Table 1, model 3), and are localized in the high-risk areas.
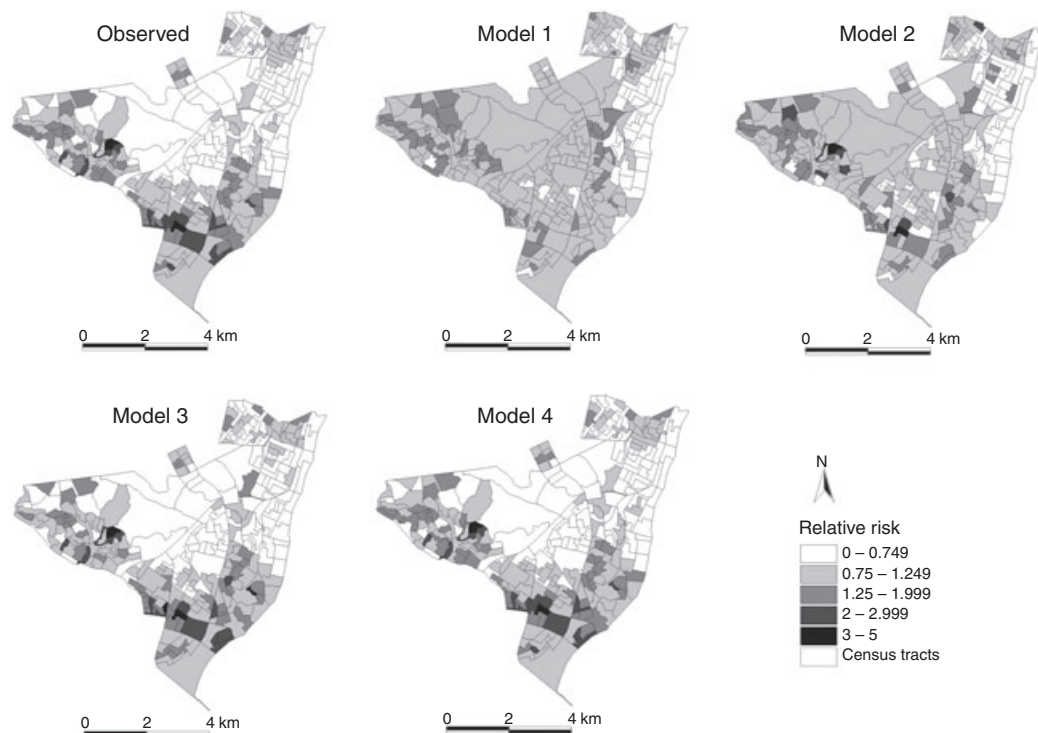

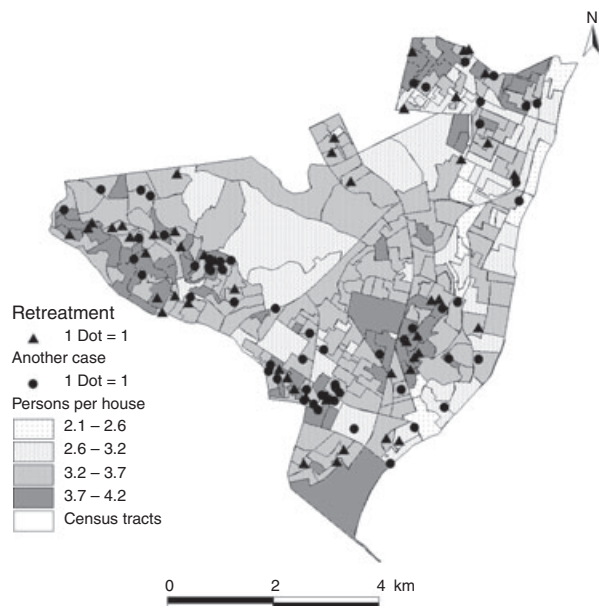**Figure 1** Observed and estimated relative risks by the four models.

Figure 2 Cases of retreatment, families with more than one case and person per house density.

Discussion

The strategy adopted to include variables in the model in related blocks allowed us to detect the importance of covariates related to the transmission of disease and the effectiveness of control measures, indirectly captured by the incidence of new cases in the same household and number of retreatments, in the proximal and distal spatial random effects model (3). First of all, their inclusion improved considerably the fit of the model. Besides, we confirmed that socio-economic variables are not, by themselves, capable of explaining the occurrence of TB: variables related to more specific dimensions (number of cases of retreatment in the district; number of households with more than one case during the period under study, per district) of the health-disease process are essential (Castellanos 1987).

Considering only the goodness-of-fit as estimated by the DIC, the models including the spatial structure were better fitted, even the model with no independent covariate. Much more precise estimates of relative risk, as confirmed by the analysis of the residuals and the best fit, were obtained when, in model 3, the variable estimating the residual of spatially-structured relative risk was introduced, confirming the hypothesis that the risk of occurrence of the disease is related to what is occurring in the neighbourhood of each spatial unit.

It is important to notice that the null spatial random effects model (4) is similar to model 3 when considering only the DIC. The DIC dropped from 1434 to 1409, but the most important changes were in the variance of random effects. The non-structured effect variance decreased by 30% and the spatial structured by almost 70%, indicating that the inclusion of fixed covariates explained part of the overdispersion of the data and the geographical structure related to TB incidence. This confirms that the spatial distribution of social deprivation is heterogeneous and clustered and that the space, as treated here, is characterized by a complex set of social processes to which different population groups are subjected, thereby defining the collective risk situation (Santos 1996).

On the other hand, random effects models *per se*, do not explain the disease incidence, neither explicit variables that can be targeted for intervening in the social processes. The number of inhabitants per household, information easily obtained in the demographic census is a risk factor for TB, and census tract with similar socio-economic characteristics tend to cluster together. Still more important is the number of cases of retreatment in each census tract. Together, these variables show the role of different social processes, the first one indicating a structural dimension of quality of life, while the second one points to access to and quality of provision of public health care.

A limitation of this approach is that most TB control programmes (especially, in high burden and low income countries) do not have access to high-level statistical programmes, nor do they often have the expertise to perform such analyses. On the other hand, it offers the opportunity of integration between the university and the services enabling further collaboration and discussion on strategies for control. Our results pose two different alternatives to deal with controlling TB in this area. The use of variables related to the performance of the control programme and, more specifically, the association with retreatment points to the measures to be implemented at the individual level, while the mapping of risk indicates the need to reorganize health services targeting subgroups of the population with a higher risk of the disease. In addition, the association either with income and household crowding or with the spatial organization, highlights the links between social deprivation and TB in this setting.

The adoption of a full hierarchical Bayesian model to estimate the relative risk of the occurrence of TB was an adequate method for this study. As 'all models are wrong, but some are useful', the possibility to estimate in the same framework both the contribution of covariates at ecological level and the spatial pattern as well should be encouraged in epidemiology, in spite of the inherent difficulties of this approach (Hampel 1987). The

formalization of the previous knowledge in the model, in the form of the *a priori* probability distribution of a hypothesis can improve the usefulness of the models. Besides, identifying the social processes that lead to the situations of collective risk are necessary to rational planning of interventions, and may help with establishing Epidemiological Surveillance Systems on a territorial basis.

References

Aber VR, Allen BW, Mitchison DA, Ayuma P, Edwards EA & Keyes AB (1980) Quality control in tuberculosis bacteriology. 1. Laboratory studies on isolated positive cultures and the efficiency of direct smear examination. *Tubercle* **61**, 123–133.

Bailey TC (2001) Spatial statistical methods in health. *Cadernos de Saúde Pública* **17**, 1083–1098.

Bailey TC, Carvalho MS, Lapa TM, Souza WV & Brewer MJ (2005) Modeling of under-detection of cases in disease surveillance. *Annals of Epidemiology* **15**, 335–343.

Barr RG, Diez-Roux AV, Knirsch CA & Pablos-Méndez A (2001) Neighborhood poverty and the resurgence of tuberculosis in New York city, 1984–1992. *American Journal of Public Health* **91**, 1487–1493.

Berkman LF & Kawachi I (2000) *Social Epidemiology*. Oxford University, New York.

Bernardinelli L, Clayton D & Montomoli C (1995) Bayesian estimates of disease maps: how important are priors? *Statistics in Medicine* **14**, 2433–2443.

Best N, Richardson S & Thomson A (2005) A comparison of Bayesian spatial models for disease mapping. *Statistical Methods in Medical Research* **14**, 35–59.

Camara G & Monteiro AMV (2001) Geocomputation techniques for spatial analysis: are they relevant to health data? *Cadernos de Saúde Pública* **17**, 1059–1071.

Castellanos PL (1987) Sobre el concepto de salud/enfermedad: um punto de vista epidemiológico. *Cuadernos Medicos Sociales* **42**, 15–24.

Cromley E, McLafferty S, Cromley EK & McLafferty SL (2002) *GIS and Public Health*. Guilford, New York.

Elliott P, Wakefield J, Best N & Briggs D (2001) *Spatial Epidemiology: Methods and Applications*. Oxford University Press, London.

Gilks WR, Richardson S & Spiegelhalter D (1996) *Markov Chain Monte Carlo in Practice*. Chapman and Hall, Londres.

Hampel FR (1987) Data analysis and self-similar processes. In: *Proceedings of the 46th Session of the International Statistical Institute*, book 4, Tokyo. International Statistical Institute, Voorburg, NL, 235–254.

Instituto Brasileiro de Geografia e Estatística (2001) *Censos Demográficos*. IBGE, Rio de Janeiro. Available from: http://www.ibge.gov.br/censo (accessed 5 January 2003).

Kawachi I & Berkman LF (2003) *Neighbourhoods and Health*. Oxford University Press, New York.

Krieger N, Waterman PD, Chen JT, Soobader MJ & Subramanian SV (2003). Monitoring socioeconomic inequalities in sexually transmitted infections, tuberculosis, and violence: geocoding and choice of area-based socioeconomic measures – the Public Health Disparities Geocoding Project (US). *Public Health Reports* **118**, 240–260.

Lawson A (2001) *Statistical Methods in Spatial Epidemiology*. John Wiley & Sons, Sussex.

Militino AF, Ugarte MD & Dean CB (2001) The use of mixture models for identifying highrisks in disease mapping. *Statistics in Medicine* **20**, 2035–2049.

Ricketts TC (2003) Geographic information systems and public health. *Annual Review of Public Health* **24**, 1–6.

Rose G (2001) Sick individuals and sick populations. *International Journal of Epidemiology* **30**, 427–432.

Santos M (1996) *A Natureza do Espaço*. Hucitec, São Paulo.

Singh-Manoux A, Clarke P & Marmot M (2002) Multiple measures of socio-economic position and psychosocial health: proximal and distal measures. *International Journal of Epidemiology* **31**, 1192–1199.

Souza WV, Ximenes RAA, Albuquerque MFM *et al.* (2000) The use of socioeconomic factors in mapping tuberculosis risk areas in a city of northeastern Brazil. *Revista Panamericana de Salud Pública* **8**, 403–410.

Spiegelhalter D, Thomas A, Best N & Gilks WR (1997) *BUGS: Bayesian Inference Using Gibbs Sampling*. MRC Biostatistics Unit, Cambridge.

Spiegelhalter D, Best N, Carlin BP *et al.* (2002) Bayesian measures of model complexity and fit. *Journal of Royal Statistical Society. Series B, Methodological* **64**, 539–583.

Waaler HT (2002) Tuberculosis and poverty. *International Journal of Tuberculosis and Lung Disease* **6**, 745–746.

Corresponding Author Wayner Vieira de Souza, Av. Moraes Rego s/nº, Cidade Universitária, Recife 50670-420, Brazil. Tel.: 55 81 2101 2648; Fax: 55 81 3453 3046; E-mail: wayner@cpqam.fiocruz.br

Appendix1 WinBugs code used

```
model
{
theta[1:regions] ~ car.normal(adj[], weights[], num[], tau)
theta.mean < - mean(theta[])
for(i in 1:regions){
o[i]  ~ dpois(mu[i])
log(mu[i])  < - log(e[i]) + beta0 + beta1*X1[i] + beta2*X2[i]
     + beta3*X3[i] + beta4*X4[i] + beta5*X5[i] + theta[i]+v[i]
RATEhat[i] < - 100*mu[i]/e[i]
v[i] ~  dnorm(0,phi)
}
beta0  ~  dflat()
beta1  ~ dnorm(0.0, 1.0E-5)
beta2  ~ dnorm(0.0, 1.0E-5)
beta3  ~ dnorm(0.0, 1.0E-5)
beta4  ~ dnorm(0.0, 1.0E-5)
beta5  ~ dnorm(0.0, 1.0E-5)
phi  ~ dgamma(rstar, dstar)
tau  ~ dgamma(rstar, dstar)
}
```

La tuberculose dans les zones intra urbaines: une approche bayésienne

OBJECTIF Modeler l'effet de la privation socio-économique et de quelques indicateurs associés à la transmission sur l'incidence de tuberculose au sein de zones réduites, afin de discuter le potentiel de chaque indicateur pour le ciblage des endroits où développer une action préventive.

MÉTHODES Étude spatiale écologique de l'incidence de tuberculose à Olinda, une ville dans le nord-est du Brésil durant la période 1996–2000. Nous avons utilisé trois indicateurs socio-économiques (nombre moyen d'habitants par ménage, pourcentage des chefs de famille avec moins d'un an d'enseignement conventionnel, pourcentage des chefs de familles avec un revenu mensuel inférieur au salaire minimum) et deux indicateurs associés à la transmission (nombre de cas de retraitement, nombre de ménages avec plus d'un cas durant la période de l'étude) tous calculé par région de recensement. Nous avons adopté quatre différents modèles bayésiens hiérarchiques complets pour estimer le risque relatif de l'occurrence de la tuberculose suivant la méthode de Markov chain Monte Carlo.

RÉSULTATS Le meilleur modèle inclut tous les covariables choisis et l'effet aléatoire dans l'espace structuré. Le gain dans qualité de la statistique quand la structure spatiale était incluse confirme le profil spatial en grappe de la maladie et de la pauvreté. Dans ce modèle, les covariables dans l'intervalle de crédibilité hors de zéro étaient le nombre de personnes par maison, le nombre de cas de retraitement et le nombre de ménages avec plus d'un cas (tous avec un risque relatif ≥1.8) dans chaque région de recensement.

CONCLUSIONS La possibilité d'estimer dans un même cadre la contribution à la fois des covariables au niveau écologique et le profil spatial devrait être encouragée en épidémiologie et peut aider à l'établissement de Systèmes de Surveillance Epidémiologiques sur une base territoriale, qui permet une planification rationnelle des interventions et l'amélioration de l'efficacité du Programme de Contrôle.

mots clés tuberculose, méthodes bayésiennes, analyse spatiale, privation, surveillance épidémiologique

W. V. Souza *et al.*   **Tuberculosis in intra-urban settings**

Tuberculosis en asentamientos intraurbanos: un enfoque Bayesiano

OBJETIVO   Modelar el efecto de la privación socioeconómica y algunos indicadores relacionados con la transmisión de la incidencia de la tuberculosis a nivel de áreas pequeñas, con el fin de discutir el potencial de cada indicador a la hora de escoger los lugares en los que se desarrollarían acciones preventivas.

MÉTODOS   Estudio ecológico espacial sobre la incidencia de la tuberculosis en Olinda, ciudad al noreste de Brasil, durante el periodo de 1996–2000. Se utilizaron tres indicadores socioeconómicos (número medio de habitantes por casa; porcentaje de cabezas de familia con menos de una año de educación formal; porcentaje de cabezas de familia con un ingreso mensual por debajo del salario mínimo) y dos indicadores relacionados con la transmisión (número de casos de re-tratamiento; número de casas con más de un caso durante el periodo de estudio), todos calculados por *census tract*. Adoptamos cuatro diferentes modelos bayesianos completamente jerárquicos para estimar el riesgo relativo de tuberculosis mediante Cadenas Markov de Monte Carlo.

RESULTADOS   El modelo con mejores especificaciones incluye todas las covariables escogidas y el efecto aleatorio espacialmente estructurado. La ganancia en la mejoría del ajuste estadístico cuando la estructura espacial fue incluida, confirma el patrón espacial agrupado de enfermedad y pobreza. En este modelo, las covariables dentro del intervalo de credibilidad diferente a cero fueron el número de personas por casa, el número de casos de re-tratamiento y el número de casas con más de un caso (todos con un riesgo relativo ≥1.8) en cada *census tract*.

CONCLUSIONES   En epidemiología se debería fomentar el poder estimar dentro de la misma estructura las contribuciones de las covariables a nivel ecológico y de patrones espaciales, lo cual podría ayudar en el establecimiento de sistemas de vigilancia epidemiológica a nivel territorial, que permiten la planeación racional de intervenciones y la mejora de la efectividad de programas de control.

palabras clave tuberculosis, métodos bayesianos, análisis espacial, privaciones, vigilancia epidemiológica ca