

Spatial analysis of Tuberculosis in Rio de Janeiro in the period from 2005 to 2008 and associated socioeconomic factors using micro data and global spatial regression models

Monica de Avelar Figueiredo Mafra Magalhães ¹
Roberto de Andrade Medronho ²

Abstract *The present study analyses the spatial pattern of tuberculosis (TB) from 2005 to 2008 by identifying relevant socioeconomic variables for the occurrence of the disease through spatial statistical models. This ecological study was performed in Rio de Janeiro using new cases. The census sector was used as the unit of analysis. Incidence rates were calculated, and the Local Empirical Bayesian method was used. The spatial autocorrelation was verified with Moran's Index and local indicators of spatial association (LISA). Using Spearman's test, variables with significant correlation at 5% were used in the models. In the classic multivariate regression model, the variables that fitted better to the model were proportion of head of family with an income between 1 and 2 minimum wages, proportion of illiterate people, proportion of households with people living alone and mean income of the head of family. These variables were inserted in the Spatial Lag and Spatial Error models, and the results were compared. The former exhibited the best parameters: $R^2 = 0.3215$, Log-Likelihood = -9228, Akaike Information Criterion (AIC) = 18,468 and Schwarz Bayesian Criterion (SBC) = 18,512. The statistical methods were effective in the identification of spatial patterns and in the definition of determinants of the disease providing a view of the heterogeneity in space, allowing actions aimed more at specific populations.*

Key words *Tuberculosis, Spatial analysis, Global spatial regression model*

¹ Instituto de Comunicação e Informação Científica e Tecnologia em Saúde, Fiocruz. Av. Brasil 4365, Manguinhos. 21040-360 Rio de Janeiro RJ Brasil. monica.avelar.magalhaes@gmail.com

² Instituto de Estudos em Saúde Coletiva, UFRJ. Rio de Janeiro RJ Brasil.

Introduction

Tuberculosis (TB) is caused by *Mycobacterium tuberculosis* (Koch's bacillus). The disease is a serious public health problem worldwide; it is endemic in many countries and kills approximately 1.5 million people annually¹.

TB is directly linked to poor living conditions. The probability of an individual being infected and developing the disease depends on several factors, including the socioeconomic and health conditions to which this individual is subjected². In urban centers of developing countries, social determinants such as poverty, low education level, increased population density, unhealthy housing and drug abuse constitute factors characterizing individuals who are vulnerable to the disease^{3,4}.

Social inequality, the appearance of AIDS, an aging population and large migratory movements are some of the factors highlighted by Ruffino-Neto⁵ as the main causes for the severity of the current TB situation in the world. In 2010, there were an estimated 8.5-9.2 million cases and 1.2-1.5 million deaths (including deaths from TB among HIV-positive patients). TB is the second leading cause of death from infectious diseases in the world, losing only to HIV in 2008¹.

TB remains a serious public health problem in Brazil. The highest concentration of cases is in the Southeast region, with Rio de Janeiro exhibiting the highest incidence rate. In 2010, the state of Rio de Janeiro identified 14,206 cases of TB, and the municipality of Rio de Janeiro reported the largest number of notifications: 7,664 cases. Although the incidence rate has decreased by 13% in the municipality of Rio de Janeiro in the last 9 years, in 2010, the rate was 95/100,000 inhabitants. Pulmonary TB is the predominant form of disease in this municipality^{4,6}.

The main data source for TB is the Notifiable Diseases Information System (Sistema de Informação de Agravos de Notificação- Sinan), and the notification is based on the definition of a confirmed case in the investigation and in the follow-up of cases⁷.

In the municipality of Rio de Janeiro, TB is not equally distributed in geographic space. Identifying the spatial distribution of TB and its social determinants in different areas of the municipality allows identifying more vulnerable populations and planning governmental measures more aimed at the needs of different territories. Thus, incorporating the spatial dimension in the analysis of the disease may extract addition-

al meanings than conventional analyses alone, contributing to understanding the dynamics of this disease. Therefore, the use of geoprocessing techniques, particularly Geographic Information Systems (GIS) together with Spatial Statistics, allow incorporating several variables such as location, time, socioeconomic characteristics and environmental characteristics into health studies⁸. They allow developing etiological hypotheses regarding the origin of diseases in different populations⁹. Thus, these methods seek an inferential model that includes the spatial relationships that constitute the studied phenomenon¹⁰.

A fundamental aspect in the application of these techniques is the characterization of the spatial dependency, showing how the values of the variables of interest are correlated in space. Global spatial regression models were used in some studies that associated TB and other communicable diseases with socioeconomic determinants, but those studies used larger spatial units, such as neighborhoods and administrative regions¹¹⁻¹³, rather than the census sector, for the analysis.

The present study aimed to analyze the spatial pattern of TB by identifying the relevant socioeconomic variables for the occurrence of TB by comparing the classic statistical linear regression model with models with global spatial effects.

Methodology

The present is an ecological study conducted in the municipality of Rio de Janeiro. With a population of 6,323,037, this municipality is composed of 160 neighborhoods and 763 slums (favelas), in which 22% of the population live¹⁴.

The spatial analysis analyzed area data. The census sector was chosen as the spatial unit. The TB data studied were the new cases reported to the Epidemiological Surveillance of the municipality of Rio de Janeiro through the Notifiable Diseases Information System (Sistema de Informação de Agravos de Notificação- Sinan) between the years of 2005 and 2008, provided by the Municipal Secretariat of Health and Civil Defense of Rio de Janeiro (Secretaria Municipal de Saúde e Defesa Civil do Rio de Janeiro - SMS-DC-RJ). The registries used were those whose information regarding the municipality of residence and the service were performed in Rio de Janeiro. Socioeconomic data were extracted from the Demographic Census of 2010 and served as a basis to build the indicators used in the data anal-

yses by census sector: average number of people per household (P_ANPH); average family income (AFI); proportion of heads of family with a monthly income greater than 1 minimum wage and less than 2 minimum wages (P_H_2MW); proportion of heads of family with a monthly income less than 1 minimum wage (P_H_1MW); proportion of illiterate people (P_ILLIT); proportion of households with a water supply from the general system (P_H_WS); proportion of households with a bathroom or sanitation for the exclusive use of residents and sewage system via the general sewage system or pluvial system (P_H_SS); proportion of households with garbage collection by a cleaning service (P_H_GC); proportion of households with a bathroom or sanitation for the exclusive use of residents (P_B_BS); proportion of declared white skin color (P_WSC), proportion of declared black skin color (P_BSC), proportion of declared yellow skin color (P_YSC), proportion of declared brown skin color (P_BRSC), proportion of declared indigenous skin color (P_ISC); proportion of households with people living alone (P_H_LA); and average income of the head of family (AI_H). The indicators were developed to contemplate the income, education level and housing condition dimensions to establish a proxy of the social condition of the patient.

TB cases were georeferenced by residence address using the methodology described in Magalhães et al.¹⁵. Cases in which the residence was a prison or a hospital were excluded from the study. Initially, an exploratory analysis was performed based on the incidence rate of the period for every 1,000 inhabitants, calculated by census sector.

The 314 census sectors that had no population (mostly areas occupied by massifs, ponds, green areas, etc.) or that had their data omitted by the Brazilian Institute of Geography and Statistics (Instituto Brasileiro de Geografia e Estatística - IBGE) to maintain the confidentiality of data because they had few private households were excluded from the study.

To minimize the instability of gross rates and eliminate random fluctuation, the incidence rates were smoothed using the Local Empirical Bayesian method¹⁶⁻¹⁸. A histogram showed that the data distribution was not normal, and to approximate it to a normal distribution, the neperian logarithm (Ln) transformation was used on the dependent variable.

To verify the presence of clusters, i.e., areas with their own spatial dynamics and that deserve

a detailed analysis, the Moran Global index was used.

Because this study worked with several areas, local indicators of spatial association (LISA) were used based on the neighborhood matrix generated with first order neighbors. This indicator identifies significant spatial distribution patterns and represents a decomposition of the global index¹⁹.

LISA classified the census sectors as a function of the significance level of their local index values as high/high and low/low, which indicate that they have a positive association, i.e., the location has neighbors with close values; and high/low and low/high, which indicate a negative association, i.e., the location has neighbors with different values.

After the spatial autocorrelation was confirmed, the Spearman correlation matrix was constructed, and under a statistical and epidemiological view, independent variables with significant correlation at 5% with the dependent variable and non-collinear variables, i.e., with a correlation < 0.7, were used in the analysis of classical and spatial regression.

The multivariate linear regression was applied (Ordinary Least Squares Estimation - OLS). Using the backward method and an epidemiological criterion, variables that better described the occurrence of the disease with a significant correlation at 5% were sought.

To incorporate spatial effects, the so-called models with global spatial effects that treat the spatial structure globally were applied, i.e., they suppose that the spatial correlation structure can be captured in a single parameter, which is added to the traditional regression model²⁰. To determine which model would fit to the studied variables, two alternatives were applied.

The first model used was the mixed autoregressive spatial model (Spatial Lag model), which attributes the ignored spatial autocorrelation to the response variable Y. Given that spatial dependence is considered by adding to the regression model a new term in the form of a spatial relationship for the dependent variable, then

$$Y = X\beta + \rho WY + \varepsilon$$

in which W is the spatial proximity matrix; WY expresses the spatial dependence in Y; and ρ is the autoregressive spatial coefficient¹⁶. In this model, spatial autocorrelation is incorporated as a component of the model.

The other model was the Spatial Error Model, which considers spatial effects as noise, i.e., as a

factor to be removed. This model assumes that it is not possible to model all of the characteristics of a geographic unit that may influence neighbor areas. The effects of spatial autocorrelation are associated with the error term ε , and the model is described by:

$$Y = X\beta + \varepsilon, \quad \varepsilon = \lambda W\varepsilon + \xi$$

in which is the error component with spatial effects; λ is the autoregressive coefficient; and ξ is the error component with constant and uncorrelated variance¹⁹.

When evaluating which model would better fit to the available variables, the model with highest Log-Likelihood and lowest Akaike Information Criterion (AIC) and Schwarz Bayesian criterion (SBC)^{10,20,21} was chosen.

Residuals in the Spatial Lag and Spatial Error models were analyzed using Moran's index to quantitatively verify whether the spatial autocorrelation was eliminated with the model application.

GeoDa (Arizona State University) was used to generate the models and calculate Moran's Index and LISA.

Results

There was an 11% loss in the location of cases.

In the analysis by area (Figure 1), the variation of the incidence rate after the Bayesian smoothing can be analyzed throughout the municipality of Rio de Janeiro. The highest rates appear in areas of Downtown and the surrounding neighborhoods, extending through the neighborhoods of Benfica, Mangueiras, Maré, Penha, Vila da Penha and part of the West zone.

The existence of spatial autocorrelation of the dependent variable at the level of census sectors could be observed through Moran's Global Index ($I=0.402$, $p=0.001$). The calculation of LISA classified the census sectors as a function of the significance level of their local indices values. Figure 2 shows that for every variable studied, there are areas with significant indices. Figure 2A shows at least four large agglomerates of sectors. The first agglomerate (agglomerate 1) is formed by the neighborhoods of Downtown, São Cristóvão, Tijuca, Catumbi and Glória; the second agglomerate (agglomerate 2) is formed by part of the neighborhoods of Penha and Penha Circular, which is almost an extension of the first. The third agglomerate (agglomerate 3) is formed by

Barra da Tijuca and Recreio, and the last agglomerate (agglomerate 4) is formed by portions of the neighborhoods of Pedra de Guaratiba and Barra de Guaratiba. Figure 2B shows that agglomerates 1, 2 and 4 are composed of sectors with a high incidence of TB and that agglomerate 3 comprises a group of sectors with a low incidence rate.

The Spearman correlation matrix results showed that among all independent variables studied, only the proportion of people by household did not have a significant correlation with the dependent variable. However, some variables had a very strong correlation between them: AFI, P_WSC, P_BSC and P_BRSC. Evaluating the correlations with epidemiological connotation demonstrates that the variables proportion of illiterate people (P_ILLIT), and proportion of households with garbage collection by a cleaning service (P_H_GC) had a positive correlation and therefore an opposite direction.

Based on this evaluation, the variables that did not exhibit a significant association with the dependent variable and the variables that exhibited collinearity were excluded from the OLS model. From this first result, models that better described the relationship between these variables were sought, and the chosen final model is shown in Table 1.

With this model, the R^2 determination coefficient was 0.044, the Log-Likelihood value was -10598.8, the AIC was 21207.5, and the SBC was 21243.7.

The existence of a spatial autocorrelation can be observed from the residuals of the classic regression. Moran's index of residuals was 0.3609 ($p < 0.01$). Although they have a normal distribution, they are not randomly distributed by the municipality, as shown in Figure 3.

Because of the presence of spatial autocorrelation, the Spatial Lag model used the same variables as the OLS model. In this model, the R^2 determination coefficient was 0.3215, the Log-Likelihood (value of the *likelihood function logarithm* calculated for the estimated values of the coefficients) was -9228.39, the AIC was 18468, and the SBC was 18512.2.

The residuals of this model had a normal distribution, and Moran's global index was -0.0183 ($p < 0.001$). This low value of Moran's index indicates that the inclusion of the spatial component in the model virtually eliminated the spatial autocorrelation.

For comparison purposes, the Spatial Error model was then used with the same variables of the OLS model. The R^2 determination coefficient

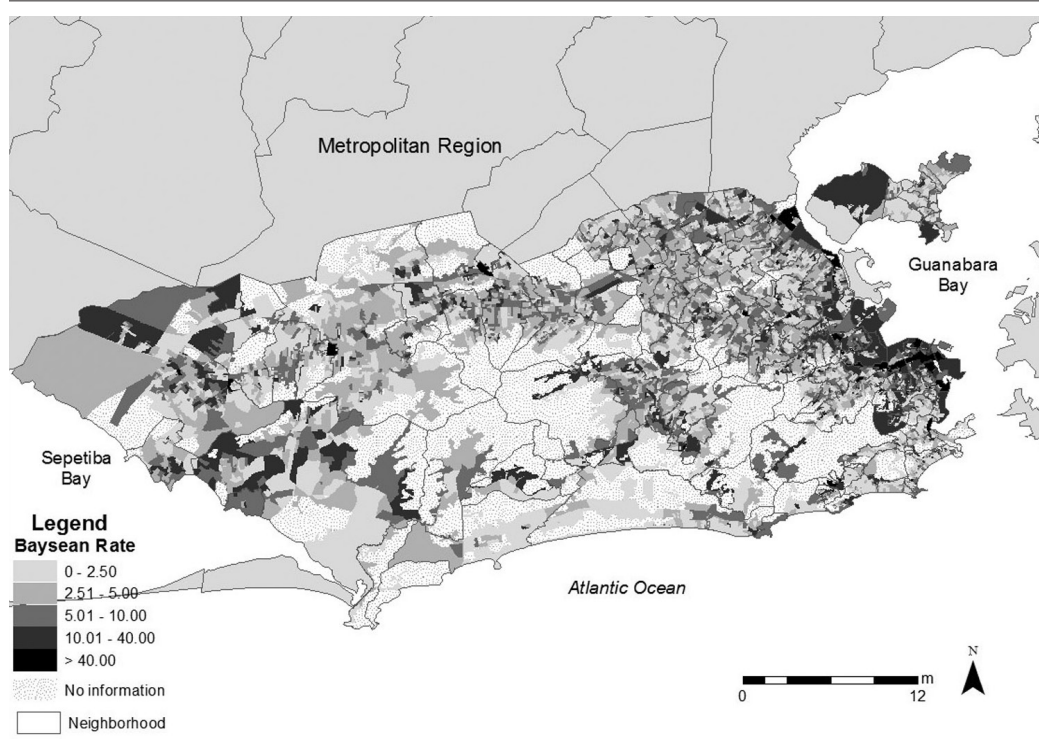


Figure 1. Map of the incidence rate of tuberculosis after the Bayesian smoothing by census sector.

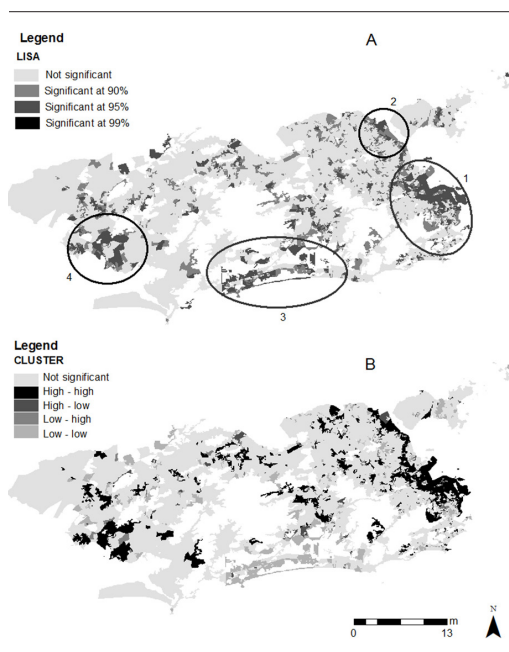


Figure 2. Map of the local indicator of spatial autocorrelation (LISA) for the dependent variable. (A) Areas with significant values; (B) LISA scattering.

was 0.319, the Log-Likelihood was -9270.72, the AIC was 18551.4, and the SBC was 18587.6.

The residuals of this model have a normal distribution, and Moran's global index was -0.0195 ($p < 0.001$). This low value of Moran's index indicates that the inclusion of the spatial component in this model also eliminated spatial autocorrelation.

Table 2 shows a summary of the indices that evaluate the quality of the models. When this spatial autocorrelation was introduced in the models, there was an improvement in the results through the Spatial Lag and Spatial Error. However, among the considered spatial regression methods, the results provided by the Spatial Lag model indicate that this model provided the best fit of the studied variables with the highest Log-Likelihood value and the lowest AIC and SBC values.

Considerations

The option to smooth TB incidence rates using the local Bayesian methods was an attempt to minimize possible distortions that may be caused by the variability provided by the calculation of

gross rates as a function of the population size at risk in the census sectors. However, a disadvantage of this method is the possible potential to induce spatial autocorrelation. According to Morais Neto et al.²², this effect may overestimate global and local autocorrelation coefficients in areas with few TB cases, in which the smoothing of census sector rates toward the averages of their neighbors is more pronounced. Although recognizing this possible bias, we chose to use this resource to correct incidence rates for two reasons. First, TB is an airborne transmitted disease with a marked influence of socioeconomic conditions. Thus, being a microarea study, when there is a census sector with a low incidence rate whose neighbors have a high rate, adjusting the rate of this sector is believed to be more coherent from the epidemiological point of view. The other reason was the large number of zeros when the gross incidence rate of tuberculosis was used. Almost half of the census sectors had no cases; therefore, they had a gross incidence rate of zero. Notably, Moran's index was also calculated with the gross rate, and the value did not differ much from the value calculated using the rate after smoothing by the Local Bayesian method.

The variables P_ILLIT and P_H_GC were correlated in the opposite direction than expected, but they were left in the analyses to evaluate how they would behave in the regression models be-

cause in the first exploratory analyses, they were important variables in the explanatory power of the model. In the final model, P_H_GC was excluded because it lost significance and P_ILLIT

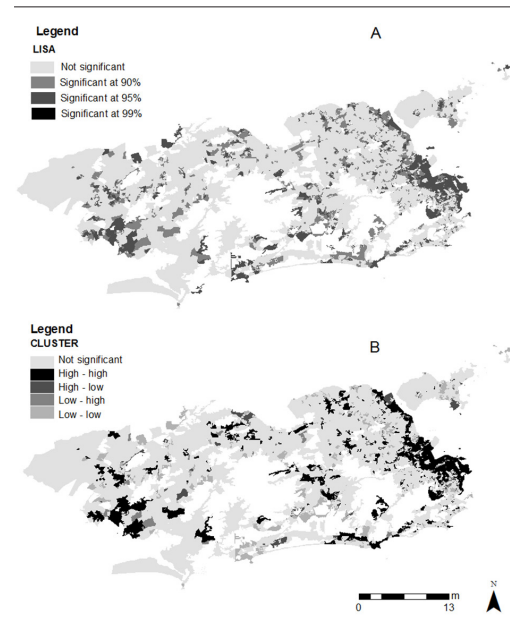


Figure 3. Map of the local indicator of spatial autocorrelation (LISA) for the residuals of the OLS model. (A) Areas with significant values; (B) LISA scattering.

Table 1. OLS model for the incidence rate log after Bayesian smoothing.

Independent variables	Coefficient	Standart error	t	p
P_R_2SM ¹	0,002958	0,000742	3,99	<0,001
P_ANALF ²	-0,017731	0,001211	-14,64	<0,001
P_DOM_MS ³	0,004201	0,000855	4,91	<0,001
RM_RESP ⁴	-0,000062	0,000004	-14,28	<0,001

¹Proportion of head of family with an income greater than 1 minimum wage and less than 2 minimum wages; ²Proportion of illiterate people; ³Proportion of households with people who live alone; 4 - Average income of the head of family.

Table 2. R^2 , Log-Likelihood, Akaike Information Criterion (AIC) and Schwarz Bayesian Criterion (SBC) indices for the three studied models.

Variable	OLS	Spatial Lag	Error Lag
R^2	0,044	0,3215	0,319
Log-Likelihood	-10598,8	-9228,39	-9270,72
AIC	21207,5	18468	18551,4
SBC	21243,7	18512,2	18587,6
Moran (I) - Residual	0,364	-0,018	-0,019

was maintained. The information regarding years of study would be very important in the present study, but this information was not collected in the 2010 Census. The information regarding education level was restricted to literate/illiterate people; therefore, the quality of the analyses from census data was reduced because together with income, this variable was a good estimator of the life conditions of the population.

The low value of the OLS determination coefficient may indicate that other variables can be related to the incidence of the disease. When comparing the three models, we can notice that those that considered spatial dependence (Spatial Lag and Spatial Error) had a better performance than the classical model (OLS). This fact can be explained by proving the spatial dependence of the dependent variable, demonstrated by the observed value of Moran's Global index and by scattering maps of Moran's Local index. Another important parameter for the evaluation of the improvement of spatial models when compared to the classical model can be observed by a significant gain in the power of explanation of variables through the value of R^2 .

Between the two spatial regression models, the Spatial Lag model exhibited the best values for Log-Likelihood, AIC and SBC.

The analysis of Moran's index values of the spatial models' residuals captured the existing spatial dependence given that values went from 0.3609 for the classical model to -0.0183 and -0.0195 for the Spatial Lag and Spatial Error models, respectively.

The best fitted model had two variables that were associated with income, P_H_2M and W_andAI_H , but P_H_2M was considered to have an indirect relationship because it is a proportion and represents a well-defined part of the population. Considering that the variable P_H_1M was also inserted in the initial analyses but did not remain in the final model because it was not significant in that model, one can indicate that the variable P_H_2M is an important cut-off point in the study of TB. Some studies^{11,23} indicated a possible association between TB infection and being part of the middle/lower class, particularly when we observe the TB infection associated with HIV and the resistant forms¹¹. A study developed in Olinda²³ showed that the relationship between the incidence rate of TB was not directly associated with the low-income population, particularly when there are conditions that favor the propagation of the disease even in places in which the population has a higher income.

This idea is corroborated by the results of the present study, which showed that the neighborhoods of Downtown and its surroundings, extending through the neighborhoods of Benfica, Manguinhos, Maré, Penha, Vila da Penha and part of the West zone (Figure 1), exhibited the highest incidence rates after the Bayesian smoothing and are areas in which the average head of family income is approximately 1,600 BRL.

The three agglomerates formed by 1) part of Downtown, São Cristóvão, Tijuca, Catumbi and Glória neighborhoods; 2) part of the Penha and Penha Circular neighborhoods; and 3) sectors of the Pedra de Guaratiba and Barra de Guaratiba neighborhoods (Figure 2) are composed of sectors with high incidence rates of TB and can be considered risk areas for the transmission of the disease.

The variable P_H_L remained in the final model. According to the literature, this variable appears inconsistent regarding TB, but a person who lives alone may have a higher risk due to non-adherence to the treatment¹¹. In a study conducted in Pelotas, Gonçalves et al.²⁴ identified that family participation is important for the adherence and permanence of patients in the treatment of TB.

The use of census sectors as a unit of analysis causes the problem of statistical instability of incidence rates because of small populations; however, it also enables working with more homogeneous populations than neighborhoods. In the study of TB in the municipality of Rio de Janeiro, the use of census sectors as a unit of analysis was very important. The analyses of the agglomerates formed by the log variable of the incidence rate after Bayesian smoothing allow the disease occurrence to be studied in greater detail. However, a limitation in the present study is related to the loss of georeferencing from addresses, described in Magalhães et al.¹⁵. The percentage of cases not located in poor areas could not be determined, but experience has shown that the vast majority of non-located cases occur in favelas. Often, the address provided by the patient is the address of the favela entrance or some reference point within the community. This causes some census tracts to receive points beyond or below expectations.

The loss in the georeferencing of cases may have removed cases referring to the most deprived population from the study described in Magalhães et al.¹⁵, and for TB in particular, it is expected that the effect on loss will be greater in the most deprived population. In addition, a study on deaths attributed to TB in the state of Rio de Janeiro per-

formed in some hospitals by Selig et al.²⁵ showed that hospitalization in the moment of notification occurred in 3,495 cases (21%), which shows problems of access, and that only 41.4% of deaths occurred in notified patients, which shows that TB in Rio de Janeiro is underreported and poorly detected. In a different study performed in Rio de Janeiro, Piller⁴ evaluated the notifying sources and determined that 26% of cases are still notified in hospitals, when these cases should have been detected and treated early by primary care. This finding reinforces the evidence that the most affected population by the disease is composed of individuals with difficulties in accessing health services

and are therefore more deprived. This population is precisely the one that lives in places of difficult localization and thus has a greater probability of being lost in the present study. Therefore, it is necessary to consider the results of the model with the caveat that some variables may have had no association with the TB rate simply because the population with the worst living conditions may have been involuntarily excluded from the analyses. However, in general, the statistical methods applied in the present study were demonstrated efficient to identify the spatial patterns of TB and to define some determinants for the occurrence of the disease.

Collaborations

MAFM Magalhães worked on conception, design, analysis, data interpretation, article writing and critical review. RA Medronho worked on conception, design, analysis and critical review.

References

1. World Health Organization. Global tuberculosis control. WHO report 2013. Geneva: WHO; 2013. [acessado 2014 mar 6]. Disponível em: http://www.who.int/tb/publications/global_report/2013/en/
2. Brasil. Ministério da Saúde (MS). *Guia de vigilância epidemiológica*. 6ª ed. Brasília: MS; 2005.
3. Buss PM, Pellegrini Filho A. A saúde e seus determinantes sociais. *Physis* 2007; 17(1):77-93.
4. Piller RVB. Epidemiologia da tuberculose. *Pulmão RJ* 2012; 21(1):4-9.
5. Ruffino-Netto A. Tuberculose: a calamidade negligenciada. *Revista da Sociedade Brasileira de Medicina Tropical* 2002; 35(1):51-58.
6. Secretária de Estado de Saúde do Rio de Janeiro, Subsecretária de Vigilância em Saúde, Sistema de Informação de Agravos de Notificação (SINAN). [acessado 2015 nov 5]. Disponível em: <http://sistemas.saude.rj.gov.br/tabnet/deftohtm.exe?sinan/tuberculose.def>
7. Pinheiro RS, Oliveira GP, Oliveira PB, Coeli CM. Melhoria da qualidade do sistema de informação para a tuberculose: uma revisão da literatura sobre o uso do linkage entre bases de dados. In: Brasil. Ministério da Saúde (MS). *Saúde Brasil 2010: uma análise da situação de saúde e de evidências selecionadas de impacto de ações de vigilância em saúde*. Brasília: MS; 2010.
8. Barcellos C, Bastos FI. Geoprocessamento, ambiente e saúde: uma união possível? *Cad Saude Publica* 1996; 12(3):389-397.
9. Medronho RA, Werneck GL. Análise de dados espaciais em saúde. In: Medronho RA, Carvalho DM, Bloch KV, Luiz RR, Werneck GL. *Epidemiologia*. Rio de Janeiro: Atheneu; 2009. p. 493-511.
10. Prado FA, Berveglieri A, Tachibana VM, Imai NN. Aplicação e análise de modelos de regressão clássica e espacial para os distritos da cidade de São Paulo. *III Simpósio Brasileiro de Ciências Geodésicas e Tecnologias da Geoinformação*; 27-30 jul. 2010; Recife - PE, p. 1-8.
11. Pereira AGL. *Distribuição espacial da tuberculose e sua correlação com variáveis socioeconômicas no município do Rio de Janeiro nos anos de 2004 a 2006* [dissertação]. Rio de Janeiro: UFRJ; 2010.
12. Sobral de Almeida A, de Andrade Medronho R, Ortiz Valencia LI. Análise espacial da dengue e o contexto socioeconômico no município do Rio de Janeiro, RJ. *Rev Saude Publica* 2009; 43(4):666-673.
13. Acosta LMW. *O mapa de Porto Alegre e a tuberculose: distribuição espacial e determinantes sociais* [dissertação]. Porto Alegre: UFRGS; 2008.
14. Instituto Brasileiro de Geografia e Estatística (IBGE). Censo demográfico, 2010. [acessado 2012 jul 2]. Disponível em <http://www.ibge.gov.br>
15. Magalhães MAFM, Pascoal VM, Medronho RA. Avaliação do dado sobre endereço no Sistema de Informação de Agravos de Notificação utilizando georreferenciamento em nível local de casos de tuberculose por dois métodos no município do Rio de Janeiro. *Cad Saúde Coletiva* 2014; 22(2):192-199.
16. Bailey TC, Gatrell AC. *Interactive spatial data analysis*. Essex: Longman Scientific and Technical; 1995.
17. Assunção RM, Barreto SM, Guerra HL, Sakurai E. Figuras de taxas epidemiológicas: uma abordagem Bayesiana. *Cad Saude Publica* 1998; 14(4):713-723.
18. Santos AE, Rodrigues AL, Lopes DL. Aplicações de estimadores bayesianos empíricos para análise espacial de taxas de mortalidade. VII Simpósio Brasileiro de Geoinformática – GEOINFO; 20-23 nov. 2005; Campos do Jordão - SP. [acessado 2011 jul 31]. Disponível em: <http://www.geoinfo.info/geoinfo2005/papers/P63.PDF>.
19. Câmara G, Carvalho MS, Cruz OG, Correa V. Análise espacial de áreas. In: Druck S, Carvalho MS, Câmara G, Monteiro AMV, organizadores. *Análise espacial de dados geográficos*. Brasília: Embrapa; 2004. p. 107-151.
20. Anselin L. Exploring Spatial Data with GeoDaTM: A Workbook. Centre for Spatially Integrated Social Science. 2005. [cited 2011 Jul 31]. Available from: <http://geodacenter.asu.edu/software/downloads>.
21. Cunha MD. *Estatística espacial na investigação epidemiológica de fatores associados à detecção de casos de hanseníase no Rio de Janeiro* [tese]. Rio de Janeiro: Fiocruz; 2012.
22. Moraes Neto OL, Barros MBA, Martelli CMT, Silva AS, Cavenaghi SM, Siqueira Júnior JB. Diferenças no padrão de ocorrência da mortalidade neonatal e pós-neonatal no Município de Goiânia, Brasil, 1992-1996: análise espacial para identificação das áreas de risco. *Cad Saude Publica* 2001; 17(5):1241-1250.
23. Souza WV, Ximenes R, Albuquerque MFM, Lapa TM, Portugal JL, Lima MLC, Martelli CMT. The use of socioeconomic factors in mapping tuberculosis risk areas in a city of northeastern Brazil. *Rev Panam Salud Publica* 2000; 8(6):403-410.
24. Gonçalves H, Costa JSD, Menezes AMB, Knauth D, Leal OF. Adesão à terapêutica da tuberculose em Pelotas, Rio Grande do Sul: na perspectiva do paciente. *Cad Saude Publica* 1999; 15(4):777-787.
25. Selig L, Belo M, Cunha AJLA, Teixeira EG, Brito R, Luna AL, Trajman A. Óbitos atribuídos à tuberculose no Estado do Rio de Janeiro. *Jornal Brasileiro de Pneumologia* 2004; 30(4):327-334.

Article submitted 19/05/2015

Approved 16/03/2016

Final version submitted 18/03/2016

