

## Spatial analysis studies of endemic diseases for health surveillance: Application of scan statistics for surveillance of tuberculosis among residents of a metropolitan municipality aged 60 years and above

Atuação da vigilância em saúde para estudos de análises espaciais de doenças endêmicas: aplicação da técnica de varredura para vigilância de tuberculose em maiores de 60 anos

Cristal Ribeiro Mesquita (<https://orcid.org/0000-0002-1321-5719>)<sup>1</sup>

Martin Johannes Enk (<https://orcid.org/0000-0002-3209-1758>)<sup>1</sup>

Ricardo José de Paula Souza e Guimarães (<https://orcid.org/0000-0002-5767-4765>)<sup>1</sup>

**Abstract** Health surveillance aims to monitor adverse health events, and to set disease prevention and control goals, especially for communicable diseases, such as tuberculosis (TB). Older people have a higher risk of TB, due to their specific characteristics, which predispose them to infectious disease. Thus, the objective of this study was to demonstrate the importance of scan statistics for detecting spatiotemporal clusters of TB. We conducted a quantitative ecological, descriptive study, with a quantitative approach, using the spatial analysis techniques, specifically scan statistics. The study was conducted in the municipality of Belém, Pará, in Brazil using data on 1,134 new cases of TB diagnosed in individuals aged  $\geq 60$  years from 2011 to 2015. The data were analyzed using SaTScan software. The analysis of the spatiotemporal dynamics of TB in the municipality showed that the high-risk areas included the most densely populated neighborhoods, highlighting the priority of these areas for disease control measures. Spatial analysis can be used to guide more effective interventions according to the characteristics of the location and the local population.

**Key words** Tuberculosis, Public Health, Spatial Analysis, Epidemiology

**Resumo** A vigilância em saúde tem como objetivo o acompanhamento dos eventos adversos à saúde, estabelecendo metas de prevenção e controle, principalmente para as doenças transmissíveis, como a tuberculose. Quando relaciona a tuberculose com a população idosa, tem-se um fator mais agravante, devido as peculiaridades dessa população que favorecem na transmissão da doença. Desse modo, objetivo do trabalho é demonstrar a importância da técnica de varredura (estatística Scan) para detectar as áreas de aglomerados espaço-temporal da TB. Trata-se de um estudo ecológico, descritivo, com abordagem quantitativa, com uso da técnica de análise espacial, em específico a técnica de varredura. O estudo ocorreu no município de Belém-PA, com 1.134 casos novos de TB em idosos no período de 2011 a 2015. Foi utilizado o programa SaTScan. A dinâmica da tuberculose no município mostrou-se que as áreas de risco coincidiram com os bairros de maior densidade populacional, evidenciando as áreas prioritárias de maior vulnerabilidade que devem servir para nortear as intervenções mais efetivas de acordo com a característica do local e da sua população.

**Palavras-chave** Tuberculose, Saúde Pública, Análise espacial, Epidemiologia

<sup>1</sup> Programa de Pós-Graduação em Epidemiologia e Vigilância em Saúde, Instituto Evandro Chagas, Secretaria de Vigilância em Saúde, Ministério da Saúde. Rodovia BR-316 KM 7 s/n, Levilândia. 67030-000 Ananindeua PA Brasil. cristalmesquita@yahoo.com.br

## Introduction

Endemic diseases, epidemics, and their impact on social conditions have long been studied and discussed in society. Measures to contain diseases and to minimize their collective impact has contributed to the develop “surveillance”. After the adoption of the Federal Constitution of 1988 (*Constituição Federal de 1988*) and the creation of the Unified Health System (*Sistema Único de Saúde - SUS*), health surveillance in Brazil became citizenship action with popular participation based on a risk approach and on social epidemiology<sup>1</sup>.

One of the fundamentals of health surveillance is epidemiological surveillance, which aims to systematically monitor adverse health events to qualify control and prevention measures, especially for communicable diseases<sup>2</sup>, such as tuberculosis, because TB is a respiratory disease.

Epidemiological surveillance advances search strategies for respiratory symptoms, not only at home but also in social spaces<sup>3</sup>. Spatial analysis, which considers the time and space of diseases, is a useful tool for epidemiological surveillance.

TB transmission in urban environments shows strong spatial determination. A specific analysis of transmission distances of *Mycobacterium tuberculosis* determined that the bacillus can be transmitted at an average distance of approximately 2,000 meters from one patient to another in the same spatial cluster, demonstrating that transmission most often does not occur at the home level<sup>4,5</sup>.

Due to the increase in life expectancy and the aging of the population, TB incidence tends to shift toward this population. TB among older individuals is a serious disease because this population has immune system specificities that favor disease transmission and hinder diagnosis, aggravating and potentiating the disease transmission chain<sup>6,7</sup>.

In 2014, the incidence rate was 34.1/100,000 inhabitants, and the incidence was 43.4 per 100,000 inhabitants in the age group of 60 years or older. In the state of Pará, the total incidence was 42.4 per 100,000 inhabitants in the general population, and 87.4 per 100,000 inhabitants in the population aged  $\geq 60$  years<sup>8</sup>. Conversely, in the capital of the state of Pará, Belém, the incidence of TB was calculated at 96.5 per 100,000 inhabitants, and the incidence in the population aged  $\geq 60$  years could not be calculated due to the unavailability of population estimates of the number of people aged  $\geq 60$  years<sup>8</sup>. As techno-

logical applications in epidemiology advance, a new branch of epidemiology emerged, spatial epidemiology, which considers the information of time and space in the analysis of epidemic diseases, describing the spatial distribution, health conditions and risk factors, thereby predicting spatiotemporal trends of diseases, such as TB, and the relationship of the disease with risk factors<sup>9</sup>.

In this context, this study aims to demonstrate the importance of scan statistics for detecting areas of spatiotemporal clusters of TB to identify risk areas and, thus, to establish prevention and control measures for this population living in risk area.

## Methods

This was a quantitative, ecological, descriptive study, conducted using spatial analysis. The study area was the municipality of Belém, capital of the state of Pará, in the northern region of Brazil. Belém has an area of 1,059,458 km<sup>2</sup>, and is divided into 71 neighborhoods, with an estimated population of 1,452,275 inhabitants, including 87,754 aged  $\geq 60$  years. The municipality has a Municipal Human Development Index (HDI-M) of 0.746 (high development)<sup>10</sup>.

The study participants consisted of 1,134 people aged  $\geq 60$  years with newly diagnosed TB recorded in the Notifiable Disease Information System (*Sistema de Informação de Agravos de Notificação - SINAN*). All new cases identified of TB in people with age over 60 years in the city of Belém-PA were utilized. Notification data were collected using standard notification forms (*Fichas de Notificação*) provided by the Municipal Health Department (*Secretaria Municipal de Saúde*) of Belém-PA (SESMA), from 2011 to 2015 (Figure 1).

The cases were georeferenced using digital maps of streets, neighborhoods, lots and house numbers provided by the Belém Metropolitan Area Development and Administration Company (*Companhia de Desenvolvimento e Administração da Área Metropolitana de Belém - CODEM*) and Google Maps.

The Georeferenced Database (*Banco de Dados Georreferenciado*) with TB cases, and vectorial files provided by CODEM and the Brazilian Institute of Geography and Statistics (*Instituto Brasileiro de Geografia e Estatística - IBGE*) were imported to an ArcGIS database to create digital maps.

Scan statistics are used to identify clusters of cases. This technique is defined by a circular

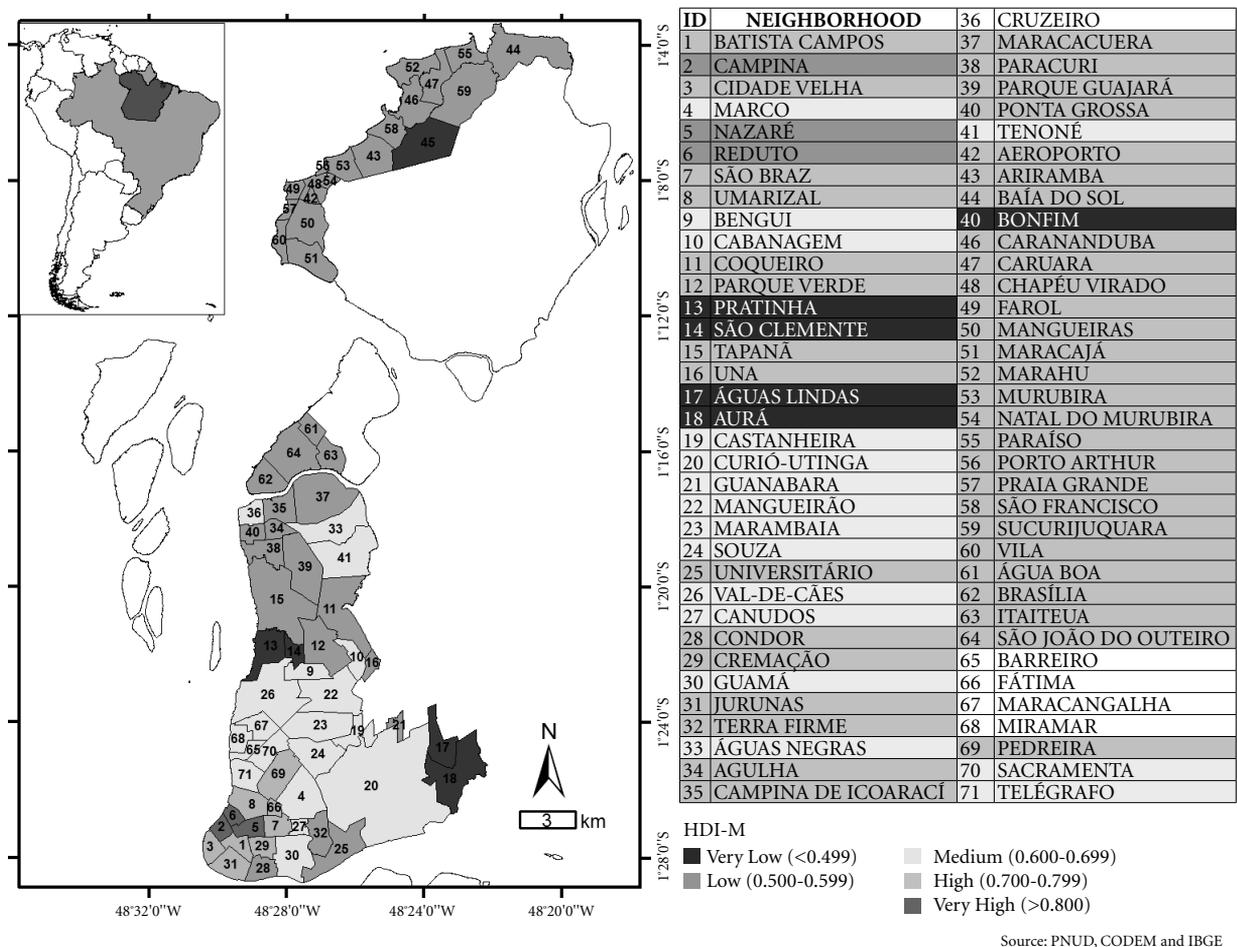


Figure 1. Spatial location of Belém neighborhoods showing the municipal human development index and the number of cases of tuberculosis in people aged ≥60 years by district.

Source: LabGeo/IEC, 2019.

scanning window that moves through the area of interest, that is, a scanning statistic where the circle can include different sets of neighboring areas<sup>11</sup>. This scanning statistic is used to detect and evaluate clusters with a temporal, spatial and spatiotemporal formation. Spatiotemporal and purely spatial analyses were performed using SaTScan software (<https://www.satscan.org/>).

A cluster of cases identified through the window may have varying scanning radiuses<sup>11</sup>. However, only some clusters are considered important. To identify these clusters, the random distribution hypothesis was tested using a discrete Poisson model (spatiotemporal) and normal model (purely spatial). The premise of the

analysis was that the risk of TB among people living inside the spatial window was higher than that among people living outside the window. The Poisson model considered the population of the study area. Statistical significance was set at  $p \leq 0.05$ .

The study was approved by the Research Ethics Committee (*Comitê de Ética em Pesquisa - CEP*) of the Evandro Chagas Institute (*Instituto Evandro Chagas - IEC*). The requirement for informed consent was waived because the study was based on a retrospective analysis of tuberculosis case notification data collected for disease surveillance purposes.

## Results

Figure 2 shows the most likely locations of clusters of TB among people aged  $\geq 60$  years in the city of Belém-PA. Figure 2A shows 3 most significant purely spatial clusters ( $p < 0.005$ ), and 2 clusters with a lower level of significance ( $p > 0.005$  and  $p < 0.010$ ), all represented by a circular scanning window. Figure 2B shows the spatiotemporal representation of cases of TB, with 2 most significant ( $p < 0.005$ ) clusters and with one less significant cluster ( $p < 0.010$ ).

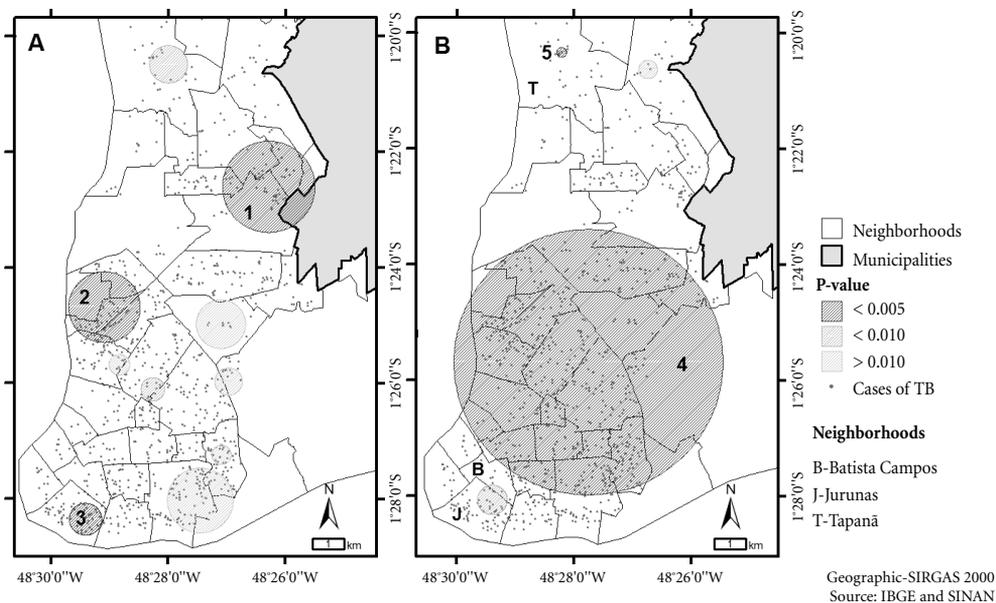
The clusters of Figure 2A are located far apart. Cluster 1 comprises 54 cases with a radius of 1.13 km, located in the neighborhoods: Manguirão, Benguí, Parque Verde, Cabanagem and Una; and has a standard deviation of 1.06 cases, and a likelihood ratio of 30.60; that is, people living in locations inside the circle are 30.60 more likely to be diagnosed with TB than those living outside the circle.

Cluster 2 comprises 33 cases within a radius of 0.51 km, located in the neighborhoods: Telégrafo, Sacramento, Miramar, Maracangalha and Barreiro; and has a standard deviation of 1.08 cases. People living in locations inside the

circle are 1.99 more likely to be diagnosed with TB than those living outside the circle.

In contrast to Clusters 1 and 2, Cluster 3, which comprises 49 cases within a radius of 1.46 km, includes only one neighborhood, Jurunas. People living in locations inside the circle are 10.75 times more likely to be diagnosed with TB than those living outside the circle. Two other, less significant clusters are shown in the neighborhoods Pedreira and Tapanã.

Figure 2B shows the spatiotemporal analysis in which the largest cluster (4), has a radius of 4.24 km, covering the entire central region of the capital of Pará. The number of cases in this cluster is 263, with 195.75 cases expected. The relative risk of the being diagnosed with TB is 1.45 times higher among those living in locations inside the circle than among those living outside the circle. The second cluster (5) 2 with  $p < 0.005$  represents a small portion of the Tapanã neighborhood, with a radius of less than 1 km. There were 6 cases of TB diagnosed among people living in the area and the expected number of cases is 0.32. The relative risk of TB among people living inside the circle is 19.08 times higher than those living outside the circle, with  $p$ -value of 0.04.



**Figure 2.** Results from the scan statistics analysis, by cylindrical scanning window, using: (A) purely spatial; and (B) spatiotemporal models.

The neighborhood Jurunas, had a more significant cluster identified using the purely spatial analysis ( $p < 0.005$ ) than the cluster identified in the spatiotemporal analysis ( $p > 0.005$  and  $p < 0.010$ ). However, in the neighborhood Tapanã, the opposite occurred, that is the cluster identified in the purely spatial analysis was less significant than the one identified using spatiotemporal analysis.

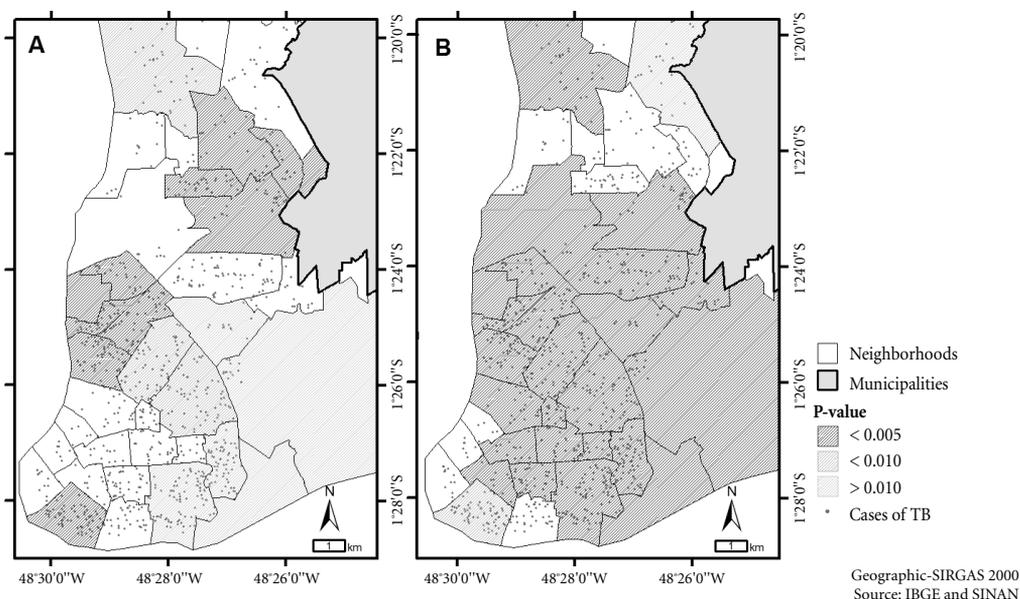
Figure 3 highlights the neighborhoods in which inhabitants have a highly ( $p < 0.005$ ), moderately ( $p > 0.005$  and  $p < 0.010$ ) and slightly ( $p < 0.010$ ) higher risk of being diagnosed with TB relative to the other, identifying the neighborhoods. Figure 3B shows the corresponding spatiotemporal analysis of TB in the study area. The number of neighborhoods with a highly significant p-value is higher than that identified in the purely spatial analysis shown in Figure 3A.

The spatiotemporal analysis showed that, when analyzed over time, almost the entire population of Belém was at risk of being diagnosed with TB due to the large spatial dispersion of cases over the time, but when analyzed by space alone, there were only 10 neighborhoods in which the risk of TB was significantly increased.

Table 1 provides the statistical details of clusters that were highly significant in the purely spatial and spatiotemporal analyses of TB case notifications in the 2011-2015 period.

The capital of Pará is heterogenous regarding its human development index (HDI). According to the 2013 United Nations Development Program (UNDP) figures<sup>12</sup>; the neighborhoods with the highest HDI are located in the central region of the capital (Batista Campos, Campina, Cidade Velha, Cremação, Fátima, Jurunas, Nazaré, Pedreira, Reduto, São Braz and Umariçal) and the neighborhoods with the lowest HDI are located in the extreme south, near Baía do Guajará, and in the far north, including Outeiro and Mosqueiro. According to the UNDP, the 10 most populous neighborhoods in Belém, were, in decreasing order: Guamá, Pedreira, Marambaia, Tapanã, Marco, Jurunas, Terra Firme, Coqueiro, Sacramento, and Telégrafo.

The first highly significant cluster identified by the purely spatial analysis was not located in any of the densely populated neighborhoods; the second was located in 2 densely populated neighborhoods (Telégrafo and Sacramento); and the third was located in the neighborhood Tapanã,



**Figure 3.** Results from the scan statistics analysis, by Belém neighborhood, in Pará (PA), using: (A) purely spatial and (B) spatiotemporal models.

**Table 1.** Statistics of clusters of TB cases with  $p < 0.005$  in the city of Belém-PA, 2011-2015.

Analysis	Number of cases	Radius (km)	LR	P-value
1-PS	54	1.13	30.90	0.001
2-PS	33	0.51	13.99	0.001
3-PS	49	1.46	10.75	0.007
4-ST	263	4.24	12.89	0.022
5-ST	6	<1.0	11.99	0.043

Abbreviations: PS: purely spatial; ST: spatiotemporal; LR: likelihood ratio.

Source: Elaborated by the authors.

which has one of the lowest HDIs of the districts within the municipality.

In the spatiotemporal analysis, the cluster with the highest significance comprised 6 of the most populous neighborhoods (Telégrafo, Pedreira, Marco, Guamá, Terra Firme and Sacramento). The neighborhood Tapanã was included in the second highly significant cluster.

## Discussion

Although methods of TB diagnosis, prevention, and treatment have been known for years, Brazil still has high incidence rates of the disease.

Belém-PA remains as a city with a high occurrence of TB, with the disease frequency only lower than Manaus-AM along the north region of Brazil. With the increase in life expectancy over the last decades, the number of people aged 60 years or older has become larger than ever<sup>13</sup>.

This group of people is more vulnerable to diseases, especially in areas with low HDI. This study's results show that TB cases in elderly people have a higher frequency in neighborhoods with lower HDI and lower population density. This shows that TB is an important disease in Belém, especially among the  $\geq 60$  years age group, because the incidence in this population is higher than in the total population. The World Health Organization (WHO) aims to reach an incidence target of 10 cases per 100,000 inhabitants, which is still far from being achieved<sup>14</sup>. The population aged  $\geq 60$  years is a key study group because health conditions among individuals in this age group can have a marked effect on life expectancy, and in addition to a higher proportion of the older population has latent *Mycobacterium tuberculosis* infection, which can be endogenously reactivated in association with chronic diseases. Timely disease diagnosis in elderly to avoid TB dissemination is also complicated.

In this context, spatial analysis makes it possible to identify risk areas, as a means to systematically monitor the factors that contribute to the transmission of TB and to support control and prevention measures<sup>2</sup>. Spatial clustering analysis is valuable in identifying geographic foci of infection in infectious disease epidemics<sup>15</sup>.

Scan statistics, developed by Kulldorff<sup>16</sup> are used to identify clusters of cases of a disease using different scanning radiuses and test for clusters using a null hypothesis that clusters of cases occur randomly in the study area as a whole. The cylindrical scanning window contains georeferenced cases and the radius increases according to the number of cases of that area<sup>16,17</sup>.

Therefore, this technique makes it possible to determine spatial locations where an event occurs more or less frequently than expected. Other studies<sup>15,17-19</sup> have used this technique to identify geographic areas where the risk of specific diseases is higher in order to flag these areas as targets for interventions in planning disease prevention and control measures.

In this study, the purely spatial analysis showed that risk areas overlapped with the neighborhoods that had the highest population density. Roza et al.<sup>20</sup> reported that a high household density, specifically overcrowded and poorly ventilated homes, facilitates TB transmission. In addition, delayed diagnosis which is common in the older population, is more common among people living in areas with a high household density. Silva et al.<sup>21</sup> reported that household density, income and education may explain differences in transmission areas according to geographic area.

Lima et al.<sup>22</sup> also conducted a study in Belém-PA, and identified neighborhoods with a higher incidence of TB. These included Val-de-Cães, Jurunas, Cremação, Fátima, Telégrafo, Pratinha, Sacramento and Umarizal. Corroborating this study, the neighborhoods Jurunas, Telegrafo and Sacramento were identified as areas at risk for

TB. It would be useful to conduct further studies in these high-risk neighborhoods aimed at identifying factors accounting for the increased incidence of TB. Importantly, the risk of TB may have been under-estimated in neighborhoods that were not included in the risk scanning windows due to case underreporting, which remains a problem in Brazil<sup>23</sup>.

This study has potential limitations due to underreporting reported by the Notification Forms of SINAN and incomplete patient's address in the form. The forms's limitations to the study should highlight, especially to health professionals, the importance of the correct notation of each case for future studies.

## Conclusion

Although TB is an ancient disease, it remains the most common notifiable infectious disease in Brazil. The clustering of cases of TB in geographic space and time, described in this study, highlights the need for new disease prevention and control

strategies, especially in high-risk geographic areas identified by the scanning technique.

Spatial analysis, complemented by other techniques, supports health surveillance in detecting priority geographic areas of higher vulnerability, and guides the choice of the most effective interventions according to the characteristics of the site and its population.

Studies using spatial analysis combined with the scanning technique have advanced the understanding of TB dynamics considerably, showing that this disease is not transmitted randomly but instead through risk clusters, suggesting a relationship between TB and the geographical space. These results corroborate previous studies on the association of this disease with socioeconomically vulnerable populations.

Further studies of the scanning technique are needed for a better understanding of TB distribution and dynamics, especially in the older population. The results of this study can be used for improved planning of health policies and health surveillance strategies.

## Collaborations

CR Mesquita, RJPS Guimarães participated in the conception, writing, review of the final version. MJ Enk participated in the review of the content until the final version of the manuscript.

## References

1. L'Abbate S, Garcia RA. Institucionalização da Vigilância em Saúde de Campinas (SP) na perspectiva da Análise Institucional sócio-histórica. *Saude Debate* 2017; 39(107):997-1007.
2. Oliveira CM, Cruz MM. Sistema de Vigilância em Saúde no Brasil: avanços e desafios. *Saude Debate* 2015; 39(104):266-267.
3. Brasil. Ministério da Saúde (MS). *Guia prático de matriciamento em saúde mental*. Brasília: MS/UFRJ; 2011.
4. Maciel ELN, Sales CMM. A vigilância epidemiológica da tuberculose no Brasil: como é possível avançar mais? *Epidemiol Serv Saude* 2016; 25(1):175-178.
5. Ribeiro FK, Pan W, Bertolde A, Vinhas SA, Peres RL, Riley L, Palaci M, Maciel EL. Genotypic and spatial analysis of *Mycobacterium tuberculosis* transmission in a high-incidence urban setting. *Clin Infect Dis* 2015; 61(5):758-766.
6. Melo CRO, Prado AM, Souza AMQ, Rolim HSE, Rosa AFS, Almeida CN. Fatores associados a tuberculose em idosos institucionalizados. *Rev Enf* 2015; 1(2):109-114.
7. Andrade SLE, Rodrigues DCS, Barrêto AJR, Oliveira AAV, Santos ARBN, Sá LD. Tuberculose em pessoas idosas: porta de entrada do sistema de saúde e o diagnóstico tardio. *Rev Enferm UERJ* 2016; 24(3):e5702.
8. Brasil. Ministério da Saúde (MS). *Panorama da tuberculose no Brasil: a mortalidade em números*. Brasília: MS; 2016.
9. Li L, Xi Y, Ren F. Spatio-Temporal Distribution Characteristics and Trajectory Similarity Analysis of Tuberculosis in Beijing, China. *Int J Environ Res Public Health* 2016; 13(3):291.
10. Instituto Brasileiro de Geografia e Estatística (IBGE). *Cidades*. Rio de Janeiro: IBGE; 2017.
11. Brasil. Ministério da Saúde (MS). *Secretaria de Vigilância em Saúde. Introdução à Estatística Espacial para a Saúde Pública. Série Capacitação e Atualização em Geoprocessamento em Saúde*. Brasília: MS; 2007.
12. Programa das Nações Unidas para o Desenvolvimento (PNUD). *Atlas do Desenvolvimento Humano no Brasil*. Brasília: PNUD/IPEA/FJP; 2013.
13. Brasil. Ministério da Saúde (MS). *Panorama da tuberculose no Brasil: a mortalidade em números*. Brasília: MS; 2016.
14. Brasil. Ministério da Saúde (MS). *Secretaria de Vigilância em Saúde. Boletim Epidemiológico*. Brasília: MS; 2016.
15. Zhao F, Cheng S, He G, Huang F, Zhang H, Xu B, Murimwa TC, Cheng J, Hu D, Wang L. Space-Time Clustering Characteristics of Tuberculosis in China, 2005-2011. *PLoS ONE* 2013; 8(21):e83605.
16. Kulldorff M. A spatial scan statistic. *Communications Statistics Theory Methods* 1997; 26(6):1481-1496.
17. Matozinhos FP, Gomes CS, Costa MA, Mendes LL, Pessoa MC, Melendez-Velasquez G. Spatial distribution of obesity in an urban Brazilian area. *Cien Saude Colet* 2015; 20(9):2779-2786.
18. Wardani DWSR, Lazuardi L, Mahendradhata Y, Kusnanto H. Clustered tuberculosis incidence in Bandar Lampung, Indonesia. *WHO South East Asia J Public Health* 2014; 3(2):179-185.
19. Manu-Yeboah D, Asarel P, Asante-Poku A, Otcherel D, Wusu O, Danso E, Forson A, Koram KA, Gagneux S. Spatio-Temporal Distribution of *Mycobacterium tuberculosis* Complex Strains in Ghana. *PLoS ONE* 2016; 11(8):e0161892.
20. Roza DL, Caccia-Bava MCGG, Martinez ED. Spatio-temporal patterns of tuberculosis incidence in Ribeirão Preto, State of São Paulo, southeast Brazil, and their relationship with social vulnerability: a Bayesian analysis. *Rev Soc Bras Med Trop* 2012; 45(5):607-615.
21. Silva MA, Oliveira CL, Teixeira Neto RG, Camargo PA. Spatial distribution of tuberculosis from 2002 to 2012 in a midsize city in Brazil. *BMC Public Health* 2016; 16(1):912.
22. Lima SS, Vallinoto ACR, Machado LFA, Ishak MOG, Ishak R. Análise espacial da tuberculose em Belém, estado do Pará. *Rev Pan-Amaz Saude* 2017; 8(2):57-65.
23. Arroy LH, Yamamura M, Protti-Zanatta ST, Fusco APB, Palha PF, Ramos ACV, Uchoa SA, Arcêncio RA. Identificação de áreas de risco para a transmissão da tuberculose no município de São Carlos, São Paulo, 2008 a 2013. *Epidemiol Serv Saude* 2017; 26(3):525-534.

---

Article submitted 15/01/2020

Approved 24/04/2020

Final version submitted 26/04/2020

---

Chief editors: Romeu Gomes, Antônio Augusto Moura da Silva