

# **Aplicação do Método de Besag & Newell Modificado para Detecção de Conglomerados Espaciais de Tuberculose no Município de João Pessoa nos anos de 2007 e 2008**

Ana Hermínia Andrade e Silva  
Ronei Marcos de Moraes

Universidade Federal da Paraíba - UFPB  
Centro de Ciências Exatas e da Natureza - CCEN  
Departamento Estatística-DE  
Laboratório de Estatística Aplicada ao Processamento de Imagens e Geoprocessamento LEAPIG  
ana\_herminia@hotmail.com  
ronei@de.ufpb.br

***Abstract** The Tuberculosis is an infect-contagious disease and its propagation is linked to the conditions of population life. The implementation of politics of health just the occurrence of the events should not be linked, but to the space where it's happen, interventions of territorial base, whose actions are gone back to risk areas meditating. This study had as objective detects space conglomerate of the tuberculosis in the municipal district of João Pessoa in the years of 2007 and 2008. A Cluster were detected in the north area of the municipal district in both years of study and in the southeast area, just a conglomerate of a neighborhood in 2007, according to the method Besag and Newell.*

***Keywords:** tuberculosis, non-parametric statistics, detect clusters.*

***Resumo** A Tuberculose é uma doença infecto-contagiosa e sua propagação está ligada às condições de vida da população. A implementação de políticas de saúde não deve apenas vincular-se a ocorrência dos eventos, mas ao espaço onde ocorrem, contemplando intervenções de base territorial, cujas ações estejam voltadas para áreas de risco. Este estudo teve como objetivo detectar conglomerados espaciais da tuberculose no município de João Pessoa nos anos de 2007 e 2008. Foram detectados conglomerados na região norte do município em ambos os anos de estudo e na região sudeste, apenas um conglomerado de um bairro em 2007, segundo o método de Besag e Newell.*

***Palavras-chave:** tuberculose, estatística não-paramétrica, detecção de conglomerados.*

## **Introdução**

Neste milênio, a tuberculose é considerada a doença infecto-contagiosa que mais mata no mundo. Um terço da população mundial está infectado e uma grande proporção dela poderá desenvolver e transmitir a doença [1]. A melhoria nas condições de vida da população [2] sugere a necessidade da implementação de políticas de saúde voltadas para populações residentes em áreas de maior risco. Estas políticas são fundamentais não apenas no monitoramento dos casos e surtos, como também no controle de fatores que possam influenciar na cura dos casos [3].

Em relação ao município de João Pessoa, eleito em 1999, um dos seis municípios prioritários para o controle da Tuberculose na Paraíba, percebe-se dificuldade de implementação da estratégia DOTS (Directly Observed Treatment Short-course) pelos serviços de atenção básica de saúde. Em 2003, as taxas de cura e de abandono do tratamento de doença foram de 61,2% e 6,4% respectivamente [4]. Na série histórica da doença, segundo os dados do SINAN (Sistema de notificação de Agravos de Notificação), o número de casos notificados de tuberculose no município sofreu uma melhora no ano 2002 em relação ao ano de 2000, porém houve um aumento anual no número de casos, no período de 2002 até 2005. Em 2005 o número de casos foi 427 (incidência de

56,1 casos/mil habitantes), em 2006, foram registrados 345 casos (46,2 casos/100 mil habitantes) e em 2007 o município registrou incidência de 52,3 casos/100 mil habitantes, valor superior a média nacional [5].

Em estudos epidemiológicos é interessante a utilização da análise espacial para detecção geográfica de doenças. Em particular os estudos de detecção de conglomerados são usados para a identificação de regiões geográficas que apresentam risco elevado em relação à ocorrência de um determinado evento, seja uma doença ou uma epidemia, em suas diversas formas [6].

Por esses motivos, o presente estudo tem como objetivo efetuar a análise da distribuição espacial da tuberculose utilizando o método de detecção de conglomerados espaciais de Besag e Newell no município de João Pessoa nos anos de 2007 e 2008.

## Metodologia

Para este estudo foram utilizados dados do Sistema Nacional de Agravos e Notificações (SINAN) referente aos casos de tuberculose residentes no município de João Pessoa, Paraíba, nos anos de 2007 e 2008. Após uma triagem do banco de dados para eliminação de informações duplicadas, obteve-se um total de 848 casos notificados que continham a informação sobre o bairro de residência. Estes dados foram obtidos junto a Secretaria de Saúde do Município de João Pessoa e permitiram a elaboração de um banco de dados.

Foi aplicada uma análise espacial sobre o número de casos de tuberculose de residentes no município de João Pessoa que informaram seu bairro de residência. Para tanto foi utilizado o método geométrico [7], iniciando então pela taxa de crescimento populacional:

$$r = \frac{\frac{Pop(t + \Delta t)}{P(t)}}{\Delta t - 1} \quad (1)$$

onde  $P(t)$  é a população no tempo  $t$ ;  $Pop(t+\Delta t)$  é a população no tempo  $t+\Delta t$ ;  $t$  é a data base;  $\Delta t$  é o intervalo entre a data base e a data a ser estimada.

A partir da taxa de crescimento obtida projetou-se a população utilizando a seguinte expressão:

$$Pop(t+\xi) = P(t)(1+r)t \quad (2)$$

onde  $Pop(t+\xi)$  corresponde à população projetada para o tempo  $t+\xi$ .

Para verificar a normalidade da distribuição dos dados de tuberculose foi aplicado o teste não-paramétrico de Lilliefors [8]. Este teste é uma derivação do teste de aderência de Kolmogorov-Smirnov que pondera os valores centrais e extremos da distribuição. Tal teste diz respeito ao grau de concordância entre a distribuição de um conjunto de valores amostrais com a estatística do teste ( $p$ -valor) calculada e com base no nível de significância adotado, pode-se aceitar ou rejeitar a hipótese nula de que os dados seguem uma distribuição normal.

Para a detecção de conglomerados espaciais foi utilizado o método não paramétrico de Besag e Newell, que não necessita do conhecimento da distribuição dos dados, pois perante a não normalidade do conjunto uma metodologia não paramétrica é a mais indicada. Besag e Newell propuseram uma saída visual de conglomerados verossímeis de formato circular que tomam por base círculos centrados sobre os centróides de cada sub-região [9]. É determinado então o raio necessário para que o círculo contenha pelo menos  $k$  casos em seu interior. O procedimento é iniciado com um círculo de raio igual a zero. Se este contém  $k$  ou mais casos, o processo é interrompido; caso contrário, o raio é aumentado até incluir o centróide mais próximo. Logo, são adicionados os respectivos casos e a população exposta ao risco. Deste modo, os raios são definidos

de tal forma a incluir novos centróides quando necessário, sendo executado este procedimento até que haja pelo menos  $k$  casos inseridos no círculo [10].

Seja  $C$  o número total de casos em toda região de estudo e  $Y$  a população total exposta ao risco na região. Sejam ainda  $C_j(i)$  e  $M_j(i)$  o número de casos e a população acumulada das  $j$  áreas mais próximas ao centróide  $i$ . A estatística do teste baseia-se na variável aleatória  $L$ , definida como o mínimo de áreas próximas necessárias para que se obtenha um número  $k$  de casos mais próximos do centróide [10]. Temos então:

$$L = \min\{j : C_{j(i)} \geq k\} \quad (3)$$

Para cada centróide é necessário verificar a existência de aglomeração espacial. Portanto, a partir do valor  $l$  observado para  $L$ , o nível de significância do teste é definido por  $P(L \leq l)$ , que busca testar a hipótese nula, que é a hipótese de ausência de aglomeração espacial. A significância é dada por:

$$p_k(i) = P(L_i \leq l_i) = 1 - \sum_{j=1}^{k-1} \frac{(M_{j(i)}C/M)^j}{j!} \exp(-M_{j(i)}C/M) \quad (4)$$

onde  $M_j$  é a população observada em  $l$  áreas.

Um aglomerado é dito significativo se o  $p$ -valor obtido é menor que a significância adotada. Assim, após a obtenção de todos os círculos que englobam um número  $k$  de casos por meio do método de Besag e Newell, são desenhados no mapa apenas os círculos significativos ( $p$ -valor  $< \alpha$ ). Geralmente a escolha de  $\alpha$  é feita para permitir os muitos testes simultâneos e são considerados valores menores que a significância usual 0,05 ou 0,01[9].

Para detecção de conglomerados espaciais foi utilizado o método de Besag e Newell modificado [10]. A justificativa para esta modificação é devido ao fato do método necessitar da determinação do número de casos e do nível de significância para busca de conglomerados, enquanto outros métodos não têm essa dependência. Além disso, a sobreposição de círculos do método de Besag e Newell pode gerar conglomerados não circulares. Assim a modificação, denominada de CABN consiste em eliminar a dependência de  $k$  e da significância no método de Besag e Newell, utilizando a seguinte estatística:

$$T = \min_k \{ \min_{z_k \in Z_k} [p_k(i)] \} \quad (5)$$

onde  $Z_k$  é o conjunto de todas as regiões circulares;  $z_k$  centradas nos centróides e contendo  $k$  casos próximos;  $p_k(i)$  é dado pela equação 2.

**Observação:** Neste artigo a variação do parâmetro  $k$  foi entre dois casos e o máximo observado em um único bairro naquele mês.

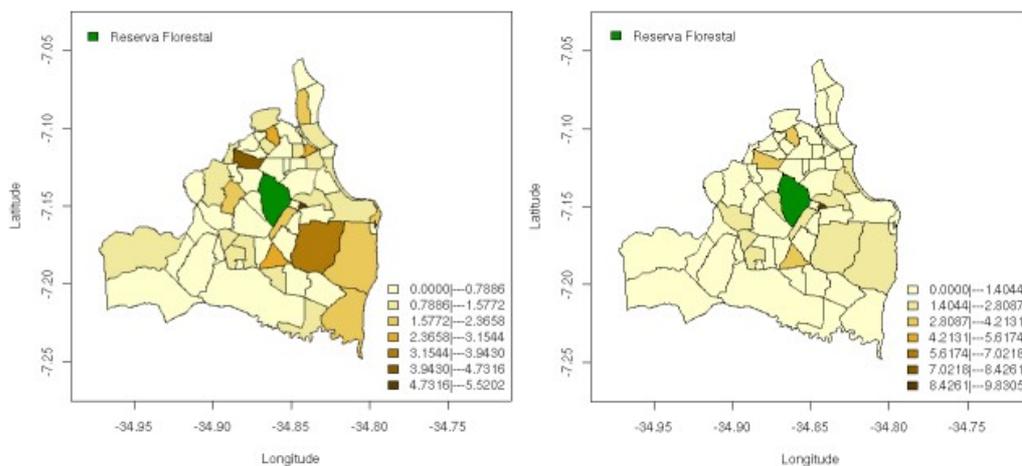
Para a geração dos resultados foi utilizado o *software* R [11]. O *software* R consiste em um sistema de análises estatísticas composto por um amplo e variado conjunto de métodos estatísticos tradicionais e modernos, sendo este um software livre e obtido gratuitamente. A versão utilizada neste estudo foi a 2.8.1 e para a geração dos resultados foram utilizadas as bibliotecas “Survival” e “DCluster”.

Para o estudo da tuberculose, fez-se necessária a projeção da população por bairro para os anos de estudo. Com base na Contagem Populacional de 2007 e no Censo Demográfico de 2000, dados obtidos no IBGE, calculou-se a taxa de crescimento e a população total para o mês correspondente a metade do período anual. Em seguida, calculou-se a proporção populacional para

cada bairro, baseando-se nos dados do censo de 2000 e considerando o crescimento populacional constante, obtendo assim a população de cada um dos bairros para os anos do período estudado.

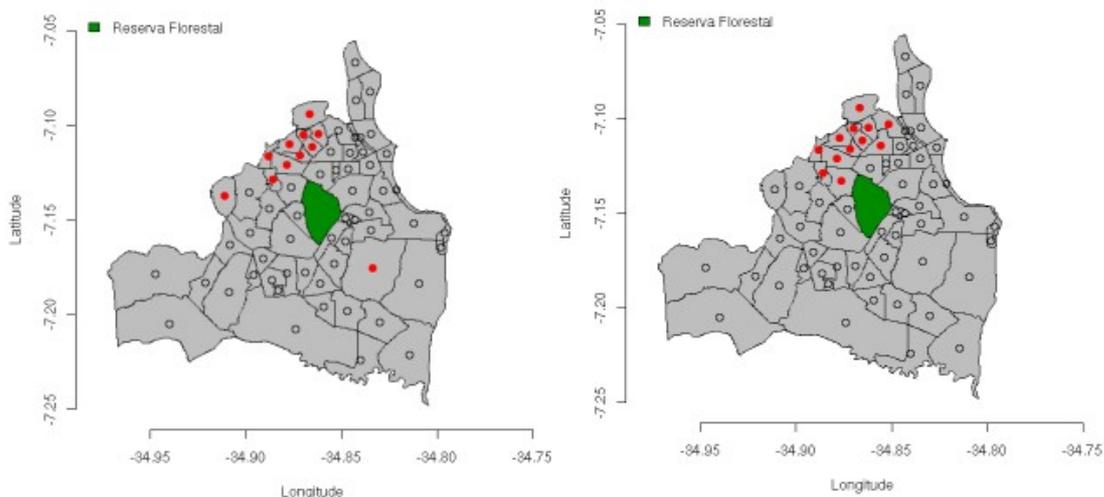
Para verificar a aderência dos dados de tuberculose sobre uma distribuição normal foi aplicado o teste de Lilliefors nos anos do período estudado. Obteve-se um p-valor menor que 0,0001, mostrando assim a não normalidade do conjunto.

Para melhor visualização dos resultados foram gerados mapas coropléticos do risco relativo da tuberculose nos anos de estudo. O cálculo do risco relativo consiste no quociente entre a incidência da da tuberculose no bairro  $i$  e a incidência total do município. Pode-se observar na Figura 1 que o risco no ano de 2007 foi acima de zero em todas as regiões do município exceto na região sudeste, com bairros com maior risco nas regiões norte e sudeste. Já no ano de 2008 o maior risco foi quase o dobro maior risco do ano anterior, porém apenas um bairro localizado na região sudeste teve um risco tão elevado em relação ao restante do município. Outra região com maiores valores de risco em relação ao restante do município nesse ano foi a região norte.



**Figura 1. Risco relativo da tuberculose nos anos de 2007 e 2008**

Para a aplicação do método Besag e Newell utilizou-se como limite superior para definir um conglomerado o valor máximo de casos do bairro do conglomerado que está sendo investigado. O nível de significância adotado foi de 5%. Pode-se observar na Figura 2 que o método detectou conglomerados significativos na região norte para ambos os anos de estudo e para o ano de 2007 detectou um conglomerado na região sudeste de apenas um bairro. Observe que os resultados do método se assemelham ao resultado esperado de acordo com os mapas de risco relativo (Figura 1).



**Figura 2. Método Besag e Newell para os anos de 2007 e 2008**

## Conclusões

Os dados utilizados nesse trabalho são referentes aos casos de tuberculose nos anos de 2007 e 2008 no município de João Pessoa, Paraíba. Foi aplicado um teste de normalidade no conjunto detectando assim a sua não normalidade. Logo foi utilizada uma metodologia não paramétrica para a detecção de conglomerados espaciais.

O método de Besag e Newell detectou conglomerados menores na região norte para ambos os anos e apenas um conglomerado de um bairro na região sudeste, para o ano de 2007. Conclui-se então, que os resultados do método de Besag e Newell são coerentes com os mapas de risco relativo, verificando ainda que na região sudeste no ano de 2007 o conglomerado detectado foi de apenas um bairro, sendo que, de acordo com o mapa de risco relativo, espera-se um conglomerado de três bairros nesta região.

## Referências

- [1] Kritski, AL, Villa, TS, Trajman, A, Silva, JRL, Medronho, RA, Ruffino Neto, A. Duas décadas de pesquisa em tuberculose no Brasil: estado da arte das publicações científicas. Rev. Saúde Pública 2007; 41(1): 9-14.
- [2] Bierrenbach, AL, Duarte, EC, Gomes, ABF, Souza, MFM. Tendência da mortalidade por tuberculose no Brasil, 1980 a 2004. Rev. Saúde Pública 2007; 41 (1):15-23.
- [3] Mota, FF, Vieira-da-Silva, LM, Paim, JJ, Costa, MCN. Distribuição espacial da mortalidade por tuberculose em Salvador, Bahia, Brasil. Cad. Saúde Pública 2003; 19(4): 915-922.
- [4] Marcolino ABL, Nogueira JA, Ruffino Neto A, Moraes RM, Sá LD, Villa TCS, Rolin FJ. Avaliação do acesso às ações de controle da tuberculose no contexto das equipes de saúde da família de Bayeux – PB. Rev. Brasileira de Epidemiologia 2009; 12(2) :144-57.
- [5] Alencar NPDD. Tuberculose no idoso: visão sobre a doença, a assistência e suas implicações no contexto familiar, João Pessoa: UFPB 2009. Dissertação (mestrado). Programa de Pós-Graduação em Enfermagem, Universidade Federal da Paraíba – João Pessoa, 2009.

- [6] Hijjar, MA, Oliveira, MJPR, Teixeira, GM. A tuberculose no Brasil e no mundo. Boletim de Pneumologia Sanitária 2001; 9(2): 9-16.
- [7] Spiegelman, M. (1968). "Introduction to Demography". London: Harvard University Press.
- [8] Lehmann EL, D'Abrera HJM. Nonparametrics Statistical Methods Base don Ranks. San Francisco: Holden-Day; 1975.
- [9] Carvalho, D.M.; Moraes, R.M. (2007). "Análise de conglomerados de homicídios por arma branca e arma de fogo na cidade de João Pessoa entre 2002 e 2005". In: VII Encontro Regional de Matemática Aplicada e Computacional. Natal. VII Encontro Regional de Matemática Aplicada e Computacional.
- [10] Costa, M. A.; Assunção, R. M. "A Fair Comparison Between the Spatial Scan and the Besag - Newell disease Clustering Tests. Environmental and Ecological Statistics 2005, 12: 301-319.
- [11] R Development Core Team. R: A language and enviroment for statistical computing. R Foundation for Statistical Computing. <http://www.R-project.org/> (acesso em 20/Maio /2010).