



Boletim de Ciências Geodésicas

ISSN: 1413-4853

bcg\_editor@ufpr.br

Universidade Federal do Paraná

Brasil

ENEAS DE FIGUEIREDO LUCENA, SADRAQUE; DE MORAES, RONEI MARCOS  
DETECÇÃO DE AGRUPAMENTOS ESPAÇO-TEMPORAIS PARA IDENTIFICAÇÃO DE  
ÁREAS DE RISCO DE HOMICÍDIOS POR ARMA BRANCA EM JOÃO PESSOA, PB  
Boletim de Ciências Geodésicas, vol. 18, núm. 4, outubro-diciembre, 2012, pp. 605-623  
Universidade Federal do Paraná  
Curitiba, Brasil

Disponível em: <http://www.redalyc.org/articulo.oa?id=393937726006>

- Como citar este artigo
- Número completo
- Mais artigos
- Home da revista no Redalyc

redalyc.org

Sistema de Informação Científica

Rede de Revistas Científicas da América Latina, Caribe, Espanha e Portugal

Projeto acadêmico sem fins lucrativos desenvolvido no âmbito da iniciativa Acesso Aberto

# **DETECÇÃO DE AGRUPAMENTOS ESPAÇO-TEMPORAIS PARA IDENTIFICAÇÃO DE ÁREAS DE RISCO DE HOMICÍDIOS POR ARMA BRANCA EM JOÃO PESSOA, PB**

*Detection of space-time cluster of homicides by stabbing in the João Pessoa city,  
Paraíba state, Brazil*

SADRAQUE ENEAS DE FIGUEIREDO LUCENA<sup>1</sup>  
RONEI MARCOS DE MORAES<sup>2</sup>

<sup>1</sup> Universidade Federal de Pernambuco  
Departamento de Estatística  
Rua Hélio Ramos s/n - Cidade Universitária  
CEP: 50740-540 - Recife - PE - Brasil

<sup>2</sup> Universidade Federal da Paraíba  
Departamento de Estatística - Cidade Universitária  
CEP: 58051-900 - João Pessoa - PB - Brasil  
sadraquelucena@gmail.com, ronei@de.ufpb.br

## **RESUMO**

A estatística Scan tem sido comumente usada para identificar conglomerados espaciais de fenômenos adversos a estudos epidemiológicos. Dentre os possíveis modelos, o Poisson e o Permutação Espaço-Tempo são adequados a dados de contagem e assumem na hipótese nula que a distribuição espacial do fenômeno é aleatória. Em estudos espaço-temporais a aleatoriedade é tratada como sendo o risco de ocorrência semelhante em todas as regiões independente do período de tempo. O presente estudo teve por objetivo avaliar o comportamento da estatística Scan para os modelos probabilísticos citados, realizando determinadas alterações na restrição populacional de cada um deles. Para efeito de análise, foram utilizados dados de homicídios por arma branca ocorridos no município de João Pessoa, PB, nos anos de 2001 a 2009. A partir desses dados buscou-se identificar possíveis padrões na ocorrência de homicídios no município e analisar o comportamento espaço-temporal do fenômeno, identificando as áreas de maior risco. Após comparar os mapas do ponto de vista epidemiológico, foi verificado que a estatística Scan utilizando o modelo Poisson com restrição de 7% da população apresentou melhores resultados. O modelo indicou a presença de conglomerados ao longo dos anos nas regiões norte

e oeste do município.

**Palavras-chave:** Estatística Scan; Análise Espaço-Temporal; Homicídios por Arma branca.

### ABSTRACT

Scan statistic has been generally used to identify spatial clusters derived from adverse phenomena, as well as from epidemiological studies. Poisson and Space-Time Permutation models are appropriate for counting data, where the null hypothesis assumes randomness in a space-time distribution of a phenomenon. In space-time studies randomness is seen as the same occurrence risk in every location, regarding (expected) time period. For analysis purposes, it was considered homicide data occurred in the João Pessoa city, Paraíba state, Brazil, by stabbing between 2001 and 2009. From these data, possible patterns in homicide occurrence were investigated and therefore, space-time behavior of the phenomenon was analyzed, identifying high risk areas. After comparing maps using an epidemiological approach, it was verified that Scan statistic using Poisson model, featuring 7% of the population restriction, had the best results. The model identified the presence of frequent clusters (recidivist clusters) over the years in the northern and western regions of the city.

**Keywords:** Scan Statistic; Space-Time Analysis; White Weapon Homicides.

### 1. INTRODUÇÃO

Em estudos epidemiológicos, uma forma de compreender um fenômeno diz respeito à identificação de áreas de risco elevado em áreas específicas da região de estudo. Neste sentido, têm recebido atenção as metodologias estatísticas que fornecem o aporte teórico para identificação de padrões espaciais (ou espaço-temporais) e aglomerados de eventos espacialmente distribuídos (MOURA, 2006). A utilização desses métodos é útil para o desenvolvimento de novos instrumentos e o acompanhamento de fenômenos no espaço e no tempo. Assim, é possível fornecer instrumentos tecnológicos capazes de ampliar a prevenção do fenômeno de interesse e minimizar os danos à população exposta ao risco (PELLEGRINI, 2001).

Os métodos que buscam por regiões de risco elevado são denominados métodos de detecção de conglomerados. Um conglomerado é definido como uma região cujo risco de ocorrência de um fenômeno é alto quando comparado às demais áreas (KULLDORFF, 1997). Há na literatura diversos métodos de detecção de conglomerados espaciais e espaço-temporais, ou seja, métodos que buscam regiões de risco elevado ao longo do tempo (BESAG; NEWELL, 1991) (TANGO, 2000) (CUZICK; EDWARDS, 1990). Como exemplo, Lucena e Moraes (2009) identificaram regiões de alto risco de ocorrência do dengue em João Pessoa no norte e sudeste do município em 2004 e 2005. Tais métodos são classificados de acordo com a distribuição de probabilidade dos dados, isto é, há métodos que supõem normalidade (TANGO, 2000), outros se aplicam de acordo com a escala de

mensuração (discreta ou contínua) (KULLDORFF; NAGARWALLA, 1995) e outros ainda não fazem qualquer suposição sobre os dados (CUZICK; EDWARDS, 1990). Tais métodos têm como hipótese nula a suposição de que não há conglomerados na região de estudo ao longo do tempo, ou seja, o risco de ocorrência do evento é aleatório e proporcional ao tamanho populacional de cada localidade no decorrer do período considerado (MOURA, 2006).

Dentre os métodos de detecção de conglomerados, como o Método de Besag e Newell (BESAG; NEWELL, 1991), o Índice *I* de Moran e o Índice *C* de Geary (in ASSUNÇÃO, 2001), a estatística Scan tem recebido destaque (KULLDORFF, 1997). Além de delinear regiões críticas por meio de algoritmos computacionais gráficos, ela atribui significância à estatística de teste via simulação de Monte Carlo (KULLDORFF; NAGARWALLA, 1995). Esta metodologia pode ser aplicada a dados espaciais, temporais ou espaço-temporais, para vários modelos de probabilidade (KULLDORFF; INFORMATION MANAGEMENT SERVICES, Inc, 2009). Quando se trata de estudos com dados de contagem (por exemplo, número de homicídios por localidade) os modelos de probabilidade comumente usados são Poisson (KULLDORFF, 1997) (KULLDORFF; NAGARWALLA, 1995) e Permutação Espaço-Tempo (KULLDORFF et al., 2005). Nesta situação, ao aplicar a metodologia, convém avaliar o desempenho dos modelos para identificar qual deles apresenta melhores resultados.

Desde 1996 a Organização Mundial de Saúde (OMS) declarou que a violência é um importante problema de saúde pública, cujo aumento é considerado um fenômeno global (GAWRYSZEWSKI; KAHN; JORGE, 2005). No Brasil, o número de homicídios cresceu sistematicamente até o ano de 2003, sofrendo em seguida uma queda e pequenas oscilações. No Estado da Paraíba a taxa de homicídios registrada em 2007 foi de 43,7 para cada 100 mil habitantes. O número de homicídios no município de João Pessoa duplicou nos últimos dez anos, tornando-se a quarta capital com maior taxa de homicídios naquele ano (WAISELFISZ, 2010).

O termo “arma” se refere a todo instrumento utilizado por ser humano para ataque ou defesa. Considera-se arma branca objetos perfurantes, cortantes ou contundentes geralmente constituídos por peças em lâmina ou oblonga (GUIMARÃES, 2005). Para efeito de caracterização, vários objetos são reconhecidos como arma branca, como bengalas, cassetetes, navalhas, facas, facões, soco inglês, dentre outros.

Tendo em vista as implicações socioeconômicas dos homicídios, a identificação de áreas de risco elevado possibilita realizar uma vigilância mais efetiva, estabelecendo medidas de intervenção, através da melhor alocação de recursos humanos e capacitação de profissionais (PHEBO, 2005).

Considerando as diferentes distribuições de probabilidade que podem ser utilizadas com a estatística Scan, a definição do melhor modelo permite avaliar e monitorar adequadamente fenômenos como a violência urbana (LUCENA; MORAES, 2009) (GAWRYSZEWSKI; KAHN; JORGE, 2005). Assim sendo, este

artigo objetiva analisar o comportamento espaço-temporal dos homicídios por arma branca em João Pessoa, PB, tomando por base os resultados da estatística Scan para os modelos Poisson e Permutação Espaço-Tempo. O período considerado corresponde aos anos de 2001 a 2009, de modo a permitir avaliar a evolução das componentes espacial e temporal.

## 2. MATERIAIS E MÉTODOS

Os dados referentes às ocorrências anuais de homicídios registrados em João Pessoa, PB, foram obtidos junto à Secretaria de Estado da Segurança e da Defesa Social (SEDS). As informações foram categorizadas por homicídios causados por arma branca e agregadas por bairro de ocorrência, sendo considerados os anos de 2001 a 2009.

Os dados de população foram obtidos através do Instituto Brasileiro de Geografia e Estatística (IBGE) e correspondem ao Censo 2000 e a Contagem Populacional de 2007. Para efeito desse estudo, foram consideradas as populações por bairro de moradia dos indivíduos.

Os arquivos referentes aos mapas digitais do município foram obtidos do setor de geoprocessamento da Prefeitura Municipal de João Pessoa. Tendo em vista a divisão dos bairros ocorrida em 2006, fez-se necessário utilizar dois mapas para compreender todo o período de estudo: o primeiro corresponde aos anos de 2001 a 2005 e o segundo aos anos de 2006 a 2009. Essa divisão dos bairros ocorreu no segundo semestre de 2006, entrando em vigor e sendo distribuída em arquivo digital a partir de 2007. O primeiro mapa está dividido em 66 bairros, entretanto, após alterações realizadas em 2006, o mapa atual de João Pessoa contém 64 bairros, uma vez que dois deles foram suprimidos. Adicionalmente, Parte da Reserva Florestal do município foi dividida entre bairros adjacentes. Justifica-se a utilização nesses dois mapas pela supressão de dois bairros, bem como alterações na posição dos centroides de alguns deles.

O método Scan utilizando os modelos Poisson e Permutação Espaço-Tempo foi aplicado no *software* SaTScan<sup>TM</sup> versão 8.2.1 (KULL DORFF; INFORMATION MANAGEMENT SERVICES, Inc, 2009), que é um *software* livre e permite a análise de dados espaciais, temporais e espaço-temporais. O *software* R na sua versão 2.7.0 (R DEVELOPMENT CORE TEAM, 2008) foi utilizado para representação gráfica das saídas do SaTScan<sup>TM</sup>, facilitando a visualização e interpretação dos resultados.

### 2.1 Estatística Scan

A estatística Scan foi desenvolvida por Kulldorff e Nagarwalla (1995) e, para sua aplicação, a informação da área é associada a um único ponto dentro do polígono. Esse ponto é denominado centroide e é o centro de massa de cada localidade da região de estudo. O método então faz uma varredura por toda região em busca de áreas cuja ocorrência de um fenômeno seja significativamente mais provável.

A busca por conglomerados ocorre da seguinte forma: em cada centroide posiciona-se um círculo de raio variável e calcula-se o número de ocorrências dentro do círculo. Caso o valor observado seja maior que o esperado, a região  $z$  delimitada pelo círculo é chamada de conglomerado, caso contrário, o raio do círculo é aumentado até envolver um novo centroide. Este processo é realizado até que todos os centroides tenham sido testados. Formalmente, as hipóteses são:

$H_0$ : Não há conglomerados na região de estudo.

$H_1$ : A região  $z$  é um conglomerado.

As hipóteses são testadas por meio de um teste da razão de verossimilhanças, sendo na prática utilizada a simulação de Monte Carlo com nível de significância de 5% (MOURA, 2006). A estatística Scan também permite utilizar outras formas geométricas além da circunferência, como uma elipse (KULLDORFF et al, 2006), um retângulo (NEIL; MOORE; SABHNANI, 2005) ou outro formato arbitrário (ASSUNÇÃO; DUCZMAL, 2004) (PATIL, TALLE, 2004) (TANGO; TAKAHASHI, 2005). Adicionalmente, pode-se determinar um tamanho máximo para a figura geométrica utilizada e assim evitar conglomerados tão grandes quanto a totalidade da região em estudo. Uma estratégia é limitar a figura geométrica para que contenha no máximo um determinado percentual da população total. Não há na literatura uma formulação específica que defina o percentual adequado a cada tipo de estudo, sendo este realizado de forma empírica (LUCENA; MORAES, 2009).

Além da análise espacial, a estatística Scan permite incorporar o fator temporal no estudo. Neste caso o interesse recai sobre a identificação de conglomerados de eventos que venham a ocorrer no espaço e no tempo simultaneamente (COULSTON; RITTERS, 2003). O procedimento é análogo ao método espacial, sendo que neste caso, ao invés de círculos, são utilizados cilindros, cuja base representa o espaço geográfico e a altura representa o tempo (dia, mês, ano, etc.). A base do cilindro pode variar continuamente, para envolver uma nova unidade do mapa (bairro, setor censitário, município) e a altura do cilindro pode ser aumentada à medida que um novo período de tempo é inserido. O cilindro significativo é aquele que tem menor probabilidade de ocorrência por mero acaso.

A análise espaço-temporal pode ser realizada de duas formas: a análise retrospectiva, aplicada à análise de dados coletados em um determinado período de tempo, e a análise prospectiva, que é aplicada a dados mantidos em constante atualização. Assim, a análise retrospectiva foca no padrão dos eventos, enquanto a análise prospectiva objetiva avaliar a dinâmica do fenômeno, principalmente com relação às novas informações coletadas (YAMADA; ROGERSON; LEE, 2009). Neste trabalho foi utilizada a abordagem retrospectiva e, uma vez que foram utilizados dados de contagem de homicídios, os modelos abordados foram o modelo Poisson e o modelo Permutação Espaço-Tempo, apresentados a seguir.

#### 2.1.1 Modelo Poisson

Seja  $z$  um candidato a conglomerado,  $p(z)$  a probabilidade do fenômeno em estudo ocorrer dentro do círculo e  $q(z)$  a probabilidade do fenômeno ocorrer fora do

mesmo. No modelo Poisson são computados os raios cujos valores de  $p(z)$  e  $q(z)$  maximizam a função de verossimilhança condicionada ao total de casos observados (LUCENA; MORAES, 2009). Para o candidato  $z$  é calculada a estatística do teste da razão de verossimilhanças (GÓMEZ-RUBIO; FERRÁNDIZ-FERRAGUD; LÓPEZ, 2005):

$$\lambda = \max_{z \in Z} \frac{L(z, \hat{p}(z), \hat{q}(z))}{L_0}, \quad (1)$$

onde  $Z = \{z_1, \dots, z_k\}$  é o conjunto de todos os possíveis candidatos a conglomerados e  $\hat{p}(z)$  e  $\hat{q}(z)$  são estimativas de  $p(z)$  e  $q(z)$ .  $L_0$  é definido como

$$L_0 = \frac{C^C (M - C)^{M-C}}{M^M},$$

onde  $C$  é o total de casos observados em toda região de estudo e  $M$  é a população total.  $L(z, \hat{p}(z), \hat{q}(z))$  é definido como:

$$L(z, \hat{p}(z), \hat{q}(z)) = \frac{\exp[-p(z)n_z - q(z)(M - n_z)]}{C!} p(z)^{C_z} q(z)^{C - C_z} \prod_i c_i,$$

onde  $\exp$  representa a função exponencial;  $c_z$  e  $c_i$  ( $i, z = 1, 2, \dots, k$ ) são respectivamente o número de casos no círculo  $z$  e no círculo  $i$  e  $n_z$  é o número de indivíduos em risco no círculo  $z$ , ou seja, sob risco de serem vítimas de homicídio.

Com base na formulação apresentada, o círculo é iniciado em um único centroide e, a cada novo centroide envolvido, é calculado o valor de  $\lambda$  dado pela equação (1). Depois de calculados os valores de  $\lambda$ , registra-se aquele que possui o maior valor e a significância do teste é obtida via simulação de Monte Carlo. Este procedimento é repetido para cada centroide. O algoritmo computacional consiste dos seguintes passos (MOURA, 2006):

1. Geram-se  $B$  conjuntos de dados independentes, de modo que cada réplica possua o mesmo número  $C$  de casos que o conjunto de dados original. Os dados gerados são distribuídos ao acaso entre as  $k$  sub-regiões de acordo com a hipótese nula;
2. Calcula-se a estatística do teste da razão de verossimilhanças  $\lambda_j$  ( $j = 1, \dots, B$ ) para cada conjunto e ordenam-se seus valores;
3. Se o valor de  $\lambda$  obtido dos dados originais estiver entre os  $100\alpha\%$  maiores valores de  $\lambda_j$  simulados, rejeita-se a hipótese nula ao nível de significância  $\alpha$  (KULLDORFF, 1997).

O p-valor associado ao teste é  $1 - (R + 1)/(B + 1)$ , onde  $R$  é o número de valores de  $\lambda_j$  simulados menores que  $\lambda$ . Se a hipótese nula é rejeitada, então a região  $z$  é o conglomerado mais provável, isto é, com maior probabilidade de não ocorrer por mero acaso.

### 2.1.2 Modelo Permutação Espaço-Tempo

O modelo Permutação Espaço-Tempo possui a mesma ideia básica do modelo Poisson, entretanto três características o diferenciam: (1) este modelo só trabalha com interação espaço-tempo, ou seja, procura apenas conglomerados que ocorrem no espaço e no tempo simultaneamente; (2) tem a flexibilidade de trabalhar apenas com os casos, ou seja, não é necessário informação da população; (3) seu modelo probabilístico resulta que os casos seguem a distribuição Hipergeométrica, sob a hipótese nula (BALIEIRO, 2008).

Seja  $c_{zt}$  o número de casos observados na área  $z$  durante um tempo  $t$ . Assim, o número total de casos é dado por

$$C = \sum_z \sum_t c_{zt}.$$

O número esperado de casos para cada localização e tempo condicionado sobre as marginais observadas é definido como

$$\mu_{zt} = \frac{1}{C} \left( \sum_z c_{zt} \right) \left( \sum_t c_{zt} \right). \quad (2)$$

Para um cilindro  $A$  o número de casos  $\mu_A$  é dado pela somatória dos valores de  $\mu_{zt}$  (equação 2) que correspondem ao cilindro. O cálculo desse valor esperado é baseado na suposição de que a probabilidade condicional de haver um caso na área  $z$  em um tempo  $t_j$  é a mesma em todo tempo  $t$ .

Seja  $c_A$  o número de casos observados no cilindro  $A$ . Quando não há interação espaço-tempo,  $c_A$  tem distribuição Hipergeométrica com média  $\mu_A$ , condicionado sobre as marginais, e função de probabilidade dada por (KULLDORFF et al., 2005)

$$P(c_A) = \frac{\binom{\sum_{t \in \Omega} \sum_{z \in A} c_{zt}}{c_A} \binom{C - \sum_{t \in \Omega} \sum_{z \in A} c_{zt}}{\sum_{z \in \Omega} \sum_{t \in A} c_{zt} - c_A}}{\binom{C}{\sum_{t \in \Omega} \sum_{z \in A} c_{zt}}},$$

onde  $\Omega = A \cup A^c$ , em que  $A^c$  denota o complementar de  $A$ . Quando as duas somas no segundo termo do numerador são pequenas se comparadas a  $C$ , a distribuição de  $c_A$  é aproximadamente Poisson com média  $\mu_A$ . Assim, a Razão de Verossimilhanças Generalizada de Poisson (RVG) pode ser usada para identificar que o cilindro  $A$  é um conglomerado. O conglomerado espaço-temporal mais provável é aquele que, dentre todos os cilindros avaliados, maximiza a RVG (BALIEIRO, 2008), dada por

$$RVG = \left( \frac{c_A}{\mu_A} \right)^{c_A} \left( \frac{C - c_A}{C - \mu_A} \right)^{C - c_A}. \quad (3)$$

A significância dos candidatos a conglomerado é obtida a partir de simulações, misturando os períodos de tempo e associando-os ao conjunto inicial com a localização dos casos. Dessa forma é possível manter inalteradas as marginais espacial e temporal. Em seguida calcula-se a RVG como na equação (3) e utiliza-se



a simulação de Monte Carlo para avaliar a significância estatística de modo análogo ao apresentado no modelo de Poisson (BALIEIRO, 2008).

## 2.2 Risco relativo

Uma forma de comparar os métodos por meio da visualização cartográfica é a partir do ponto de vista epidemiológico, que visa analisar o comportamento do fenômeno sob a ótica temporal e espacial. Deste modo, a partir da análise visual dos mapas, os resultados são comparados àqueles esperados de acordo com a evolução espaço-temporal do fenômeno. Assim, os conglomerados identificados não devem ser muito distintos das regiões de ocorrência do fenômeno em estudo (ROTHMAN; GREENLAND, 1998).

Tendo em vista que os modelos descritos buscam por regiões cujo risco de ocorrência de um fenômeno é elevado em relação às demais, mapas de risco relativo foram utilizados para comparar os resultados dos modelos. O risco relativo permite comparar informações de diferentes áreas, padronizando-as e retirando assim o efeito das diferentes populações. Em suma, este indicador representa a intensidade da ocorrência de um fenômeno com relação a todas as regiões de estudo (MEDRONHO, 2004).

Seja uma região geográfica formada por áreas contíguas e disjuntas denotadas por  $A_1, A_2, \dots, A_k$ . Seja  $X$  uma variável aleatória que representa o número de ocorrências de um fenômeno, de modo que cada ocorrência  $x_i$  ( $i = 1, 2, \dots, n$ ) esteja associada a área  $A_i$ , com população  $n_i$  ( $i = 1, 2, \dots, k$ ). O risco relativo de uma área  $A_i$ , denotado por  $\theta_i$ , é o quociente entre a incidência do fenômeno observado na área  $A_i$ , e a incidência correspondente a toda área de estudo:

$$\theta_i = \frac{x_i / n_i}{\sum x_i / \sum n_i}.$$

O risco relativo foi calculado para cada ano do período considerado e foi utilizado o método geométrico para projeção da população para os anos de 2001 a 2009. Uma vez que não estavam disponíveis dados específicos para modelar o crescimento individual dos bairros, a população de cada um deles foi considerada proporcionalmente constante para a projeção.

## 3. RESULTADOS

A Secretaria de Estado da Segurança e da Defesa Social (SEDS) registrou 161 homicídios por arma branca no município de João Pessoa entre 2001 e 2009. Até o ano de 2005, a frequência anual média foi de 8,5 homicídios/ano, apresentando aumento em 2006 e 2007, que teve um pico de 42 homicídios. A partir de 2008 foi observada uma diminuição na frequência de ocorrências.

No que diz respeito aos mapas gerados, uma pequena diferença pode ser observada no período de 2001 a 2005 quando comparado com os anos de 2006 a 2009. A distinção é devido a uma reorganização dos bairros ocorrida em 2006: no mapa mais recente foram agrupados dois bairros na região oeste e outros dois na

região norte do município; alguns bairros tiveram seu território aumentado em detrimento de outros e, além disso, a parte nordeste da Reserva Florestal foi removida. Estas diferenças podem ser observadas nos mapas das Figuras 1 e 2, que apresentam os riscos de homicídio por arma branca.

### 3.1 Mapas de risco

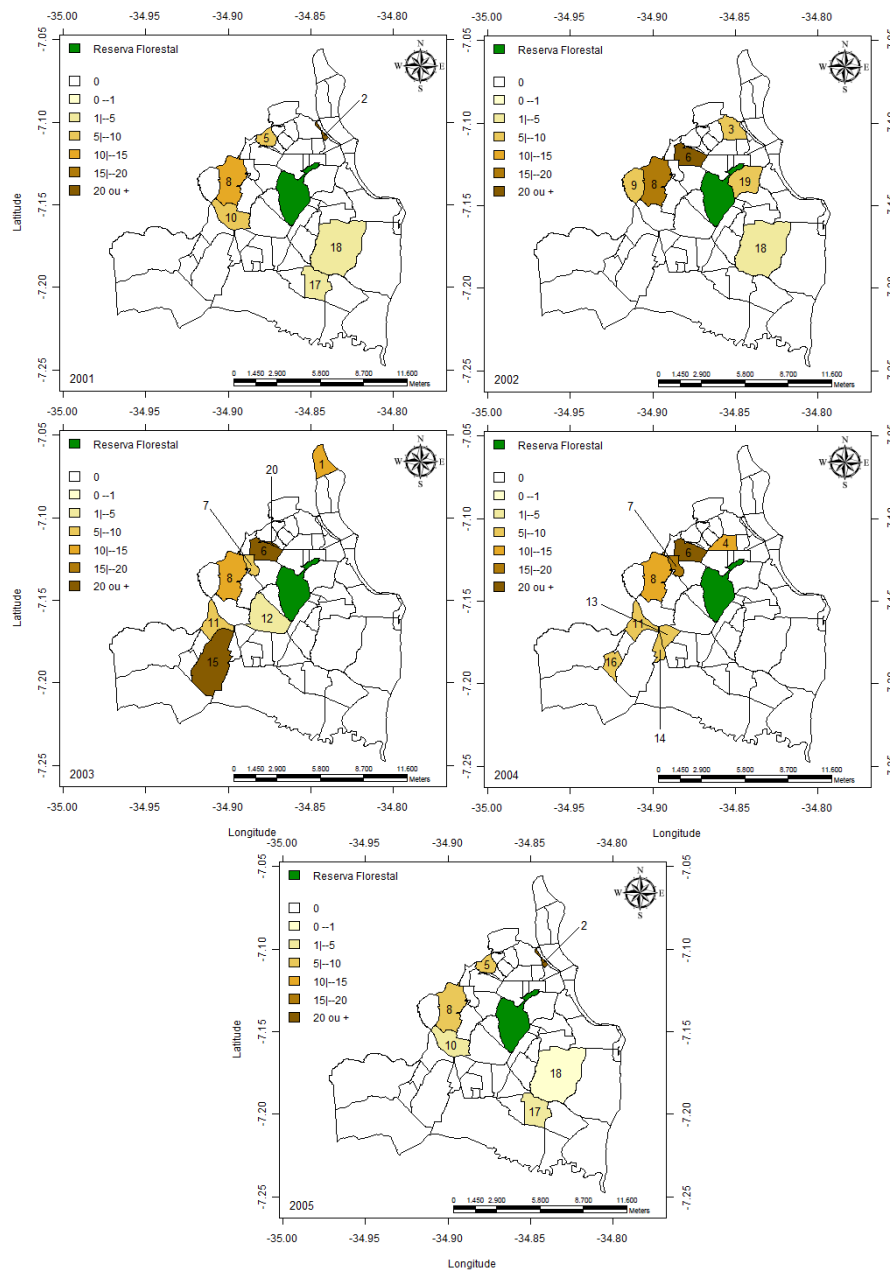
Os mapas de risco são necessários para avaliar o comportamento de fenômenos, pois a contagem pura e simples dos casos não fornece informações sobre a chance de ocorrência de homicídios nos bairros nem permite que haja comparações mais precisas entre os mesmos. Isto é, não se pode afirmar que uma região com maior número de homicídios e maior população tem risco mais elevado que outra região com frequência e população menores. Assim, o risco relativo indica as regiões com maior (ou menor) chance de ocorrência de homicídios ao padronizar as informações ponderando a incidência (quociente entre o número de ocorrências e a população) de cada bairro pela incidência do município. A descrição dos mapas apresentados é realizada com auxílio da identificação numérica dos bairros.

Como se pode observar na Figura 1, entre 2001 e 2005 os bairros que apresentaram risco de homicídio diferente de zero foram Bessa (1), São José (2), Bairro dos Ipês (3), Bairro dos Estados (4), Róger (5), Centro (6), Trinchêiras (7), Ilha do Bispo (8), Alto do Mateus (9), Oitizeiro (10), Jardim Veneza (11), Cristo (12), Ernani Sátiro (13), Costa e Silva (14), Distrito Industrial (15), Bairro das Indústrias (16), Valentina (17), Mangabeira (18), Castelo Branco (19) e Tambaú (20).

Em média, entre 2001 e 2005 não houve registro de homicídios por arma branca em 90% dos bairros, o que implica em risco zero na maior parte do município. Foi observada uma configuração espacial diferente quando comparada a distribuição dos homicídios ocorridos nos anos de 2001 e 2002. Em 2003, 71,4% dos bairros em que houve registro de homicídios não havia sido registradas ocorrências no ano anterior. A exceção ocorreu com os bairros Centro e Ilha do Bispo. No ano de 2004, 57,1% dos bairros com risco de homicídio maior que um (Centro, Trinchêiras, Ilha do Bispo e Jardim Veneza) haviam registrado casos em 2003. De modo geral, no ano de 2005, os riscos relativos observados foram menores que nos anos anteriores.

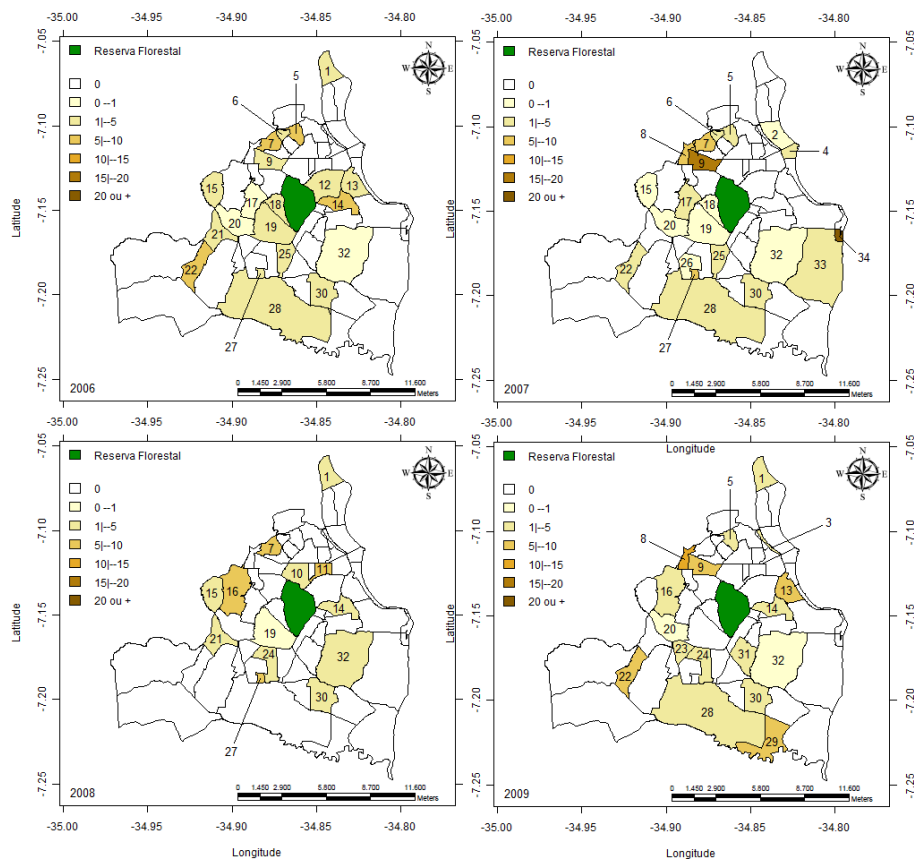
De acordo com a Figura 2, de 2006 a 2009, houve ocorrência de homicídios em 34 bairros de João Pessoa: Bessa (1), Manaíra (2), São José (3), Tambaú (4), Mandacaru (5), Padre Zé (6), Róger (7), Varadouro (8), Centro (9), Torre (10), Tambauzinho (11), Castelo branco (12), Altiplano Cabo Branco (13), Bancários (14), Alto do Mateus (15), Ilha do Bispo (16), Cruz das Armas (17), Rangel (18), Cristo (19), Oitizeiro (21), Bairro das Indústrias (22), Ernani Sátiro (23), João Paulo II (24), Ernesto Geisel (25), Funcionário (26), Grotão (27), Gramame (28), Mulumagro (29), Valentina (30), José Américo (31), Mangabeira (32), Costa do Sol (33) e Penha (34).

Figura 1 – Mapas de risco de homicídio por arma branca de 2001 a 2005.



Comparativamente aos anos anteriores, a partir de 2006 os registros de homicídio por arma branca tiveram maior abrangência no município, ocorrendo em média em 27,7% dos bairros. No ano de 2007, quinze dos 21 (71,4%) bairros com registro de ocorrências haviam apresentado risco diferente de zero no ano anterior. Em 2008 houve uma redução de 38,1% no número de bairros em que foram notificados homicídios. Em 2009 foi observada uma abrangência maior no número de bairros com registro de homicídios em relação ao ano anterior. Com exceção dos bairros em que não se evidenciaram homicídios, de modo geral, os bairros com risco menor que um (risco menor que o observado em todo município) foram Cristo Oitizeiro e Mangabeira.

Figura 2 – Mapas de risco de homicídio por arma branca de 2006 a 2009.



### 3.2 Estatística Scan espaço-temporal

A análise realizada foi do tipo retrospectiva, isto é, aquela em que o conjunto de dados é analisado sem a preocupação com informações recentes a serem inseridas. Como é necessário intervir mais fortemente em regiões com maior risco, foi realizada uma varredura circular buscando conglomerados de alto risco. Tendo em vista que na literatura não há um percentual adequado para utilização da estatística Scan, foram testados os percentuais 3%, 5%, 7% e 10% da população para identificar qual deles apresentaria melhor resultado do ponto de vista epidemiológico. Esta restrição atua da seguinte forma: os círculos utilizados pelo método têm seus tamanhos delimitados pelo número de habitantes das regiões abrangidas, isto é, cada círculo não deve conter mais habitantes do que o percentual da população definido. Quanto ao fator temporal, o modelo Permutação Espaço-Tempo busca por conglomerados que ocorram em até 50% do período, por esse motivo tal percentual também foi utilizado no modelo Poisson, para que ambos fossem tão comparáveis quanto possível. Em cada teste foram realizadas mil réplicas na simulação de Monte Carlo. Uma vez que o *software* SaTScan fornece uma saída contendo uma lista que descreve os conglomerados identificados, o *software* R foi utilizado para confecção de mapas, facilitando assim a análise dos resultados. A significância adotada foi igual a 5%.

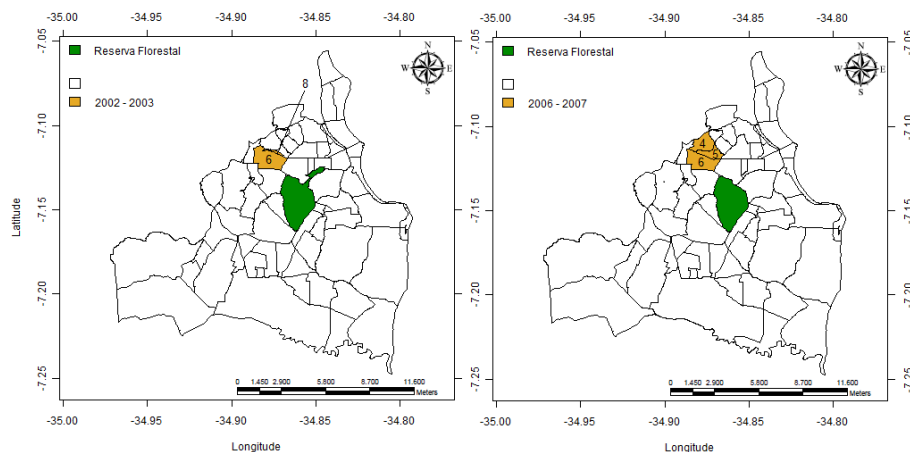
Como houve uma mudança nos limites dos bairros a partir de 2006, cada modelo foi aplicado considerando primeiramente os anos de 2001 a 2005 e em seguida, considerando de 2006 a 2009. Deste modo, os mapas correspondentes a cada período foram gerados separadamente, uma vez que a alteração dos limites dos bairros implica no deslocamento dos respectivos centroides e, conseqüentemente, na mudança das suas coordenadas geográficas.

#### 3.2.1 Modelo Poisson

Para o modelo Poisson, o primeiro percentual de população testado foi 3%, assim os conglomerados não deveriam conter mais que 3% da população do município. Na Figura 3 podem ser observados os referidos mapas com os resultados da estatística Scan, dos quais se pode concluir que:

- Para o mapa de João Pessoa que esteve em vigor de 2001 a 2005, a estatística Scan identificou um conglomerado formado pelos bairros Centro (6) e Tambiá (8), ocorrido nos anos de 2002 e 2003. O raio do círculo foi 0,59 km e o número de ocorrências observadas igual a 3, enquanto eram esperadas 0,12. O risco relativo foi igual a 27,31, a razão de verossimilhanças igual a 6,925 e p-valor igual a 0,041;
- Para o mapa de 2006 a 2009, a estatística Scan detectou um conglomerado contendo os bairros Róger (4), Tambiá (5) e Centro (6) ocorridos entre 2006 e 2007. O raio do círculo foi de 1,20 km, o risco relativo igual a 10,63, a razão de verossimilhanças igual a 21,039 e o p-valor 0,001.

Figura 3 – Conglomerados identificados usando a estatística Scan para o modelo Poisson com restrição de 3% da população para o mapa de 2001 a 2005 (esquerda) e de 2006 a 2009 (direita).



Com o percentual de população igual a 5%, percebe-se que a área identificada é maior que aquela obtida com o percentual de 3% da população. A Figura 4 apresenta os mapas com os resultados da estatística Scan para uma restrição de 5% da população:

- No mapa de 2001 a 2005, o conglomerado foi identificado nos anos de 2003 e 2004, com raio do círculo igual a 1,74 km. Os bairros envolvidos foram Róger (4), Centro (6), Varadouro (7), Trincheiras (9) e Tambiá (8). O número de homicídios observado foi igual a 6 (era esperado 0,76). O risco relativo na área identificada foi igual a 9,04, a razão de verossimilhanças igual a 7,491 e p-valor igual a 0,036;
- Em relação ao mapa de 2006 a 2009, a região identificada foi a mesma que a encontrada usando o parâmetro de restrição populacional anterior – bairros Róger (4), Tambiá (5) e Centro (6). O conglomerado ocorreu nos anos de 2006 e 2007, o raio de busca foi de 1,20 km e o número de casos observados igual a 15 (era esperado 1,59). O risco relativo foi igual a 10,63, a razão de verossimilhanças 21,039 e p-valor 0,001.

Para 7% da população, a Figura 5 apresenta os mapas com os conglomerados identificados. De acordo com os mapas, pôde-se observar que:

- A estatística Scan identificou conglomerados formado pelos bairros Padre Zé (2), Róger (4), Jardim das Acácias (5) Centro (6), Varadouro (7), Tambiá (8), Trincheiras (9) e Ilha do Bispo (10), no mapa de 2001 a 2005. O conglomerado ocorreu nos anos de 2003 e 2004, o raio do círculo foi igual a 2,47 km, o

número de ocorrências foi 8, enquanto era esperado 1,13. O risco relativo foi igual a 8,58, com razão de verossimilhanças igual a 9,435 e p-valor 0,008;

Figura 4 – Conglomerados identificados usando a estatística Scan para o modelo Poisson com restrição de 5% da população para o mapa de 2001 a 2005 (esquerda) e de 2006 a 2009 (direita).

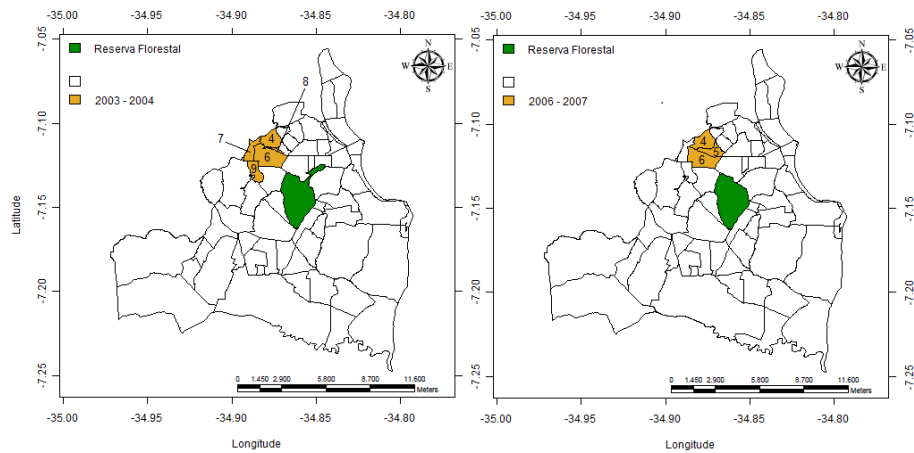
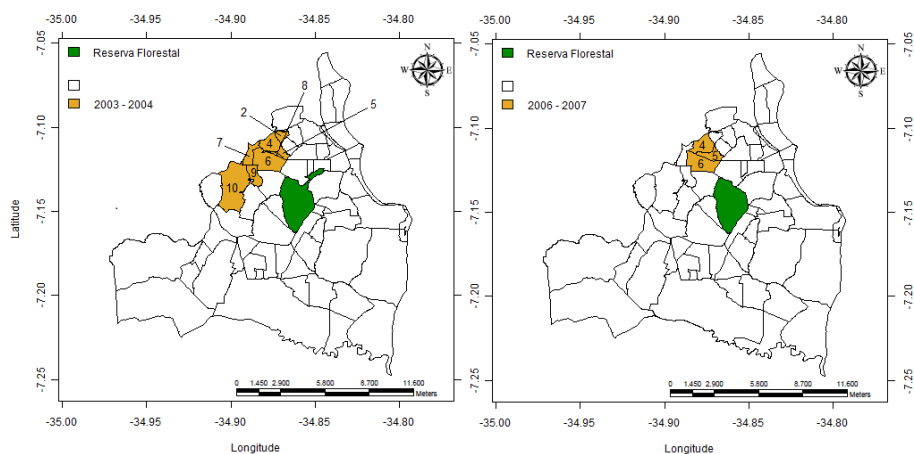


Figura 5 – Conglomerados identificados usando a estatística Scan para o modelo Poisson com restrição de 7% da população para o mapa de 2001 a 2005 (esquerda) e de 2006 a 2009 (direita).

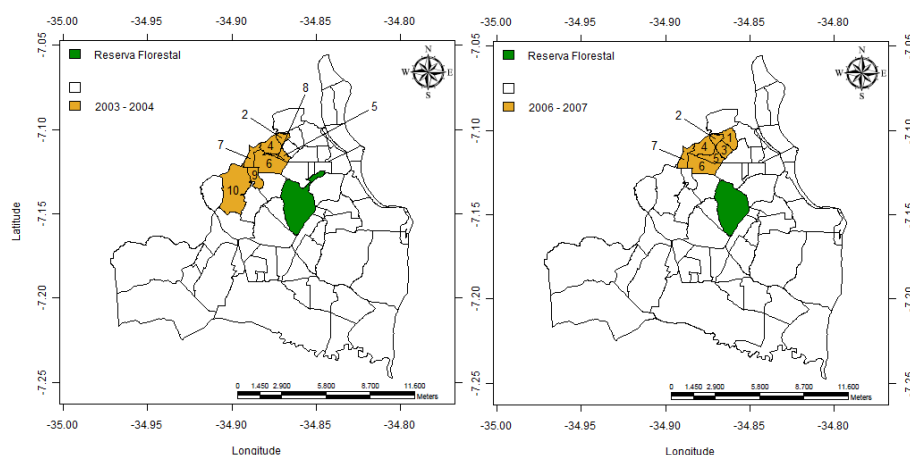


- Para o mapa de 2006 a 2009, novamente foram identificados os mesmos bairros que utilizando o modelo com restrição de 3% da população. O período correspondeu a 2006 e 2007, o raio de busca foi igual a 1,20 km, 15 ocorrências observadas e 1,59 esperadas. O risco relativo foi igual a 10,63, a razão de verossimilhanças igual a 21,039 e o p-valor 0,001.

A Figura 6 apresenta os mapas com os conglomerados identificados para uma restrição de 10% da população. Os resultados obtidos foram os seguintes:

- Para o mapa de 2001 a 2005 a estatística Scan detectou conglomerados nos bairros Padre Zé (2), Róger (4), Jardim das Acácias (5) Centro (6), Varadouro (7), Tambiá (8), Trincheiras (9) e Ilha do Bispo (10). O período identificado foi entre 2003 e 2004, o raio de busca foi de 2,47 km, o número de homicídios foi igual a oito, sendo que apenas 1,13 eram esperados. O risco relativo foi 8,58, com razão de verossimilhanças igual a 9,435 e p-valor 0,008;
- Para o mapa de 2006 a 2009, o modelo identificou um conglomerado entre os anos de 2006 e 2007, formado pelos bairros Mandacarú (1), Padre Zé (2), Treze de Maio (3), Róger (4), Tambiá (5), Centro (6) e Trincheiras (7). O raio de busca foi de 1,76 km, com 23 casos observados e 4,39 esperados. O risco relativo foi 6,24, a razão de verossimilhanças 21,047 e o p-valor 0,001.

Figura 6 – Conglomerados identificados usando a estatística Scan para o modelo Poisson com restrição de 10% da população para o mapa de 2001 a 2005 (esquerda) e de 2006 a 2009 (direita).



A seleção do percentual de população associado à identificação dos conglomerados pelo modelo Poisson foi realizada do ponto de vista epidemiológico. De modo geral, os modelos com restrição de 3% e 5% da população identificaram



conglomerados apenas nos bairros que apresentaram alto risco. Por outro lado, o modelo com restrição de 7% da população além de identificar conglomerados em bairros cujo risco foi elevado, os círculos envolveram também alguns bairros adjacentes com riscos mais baixos. Adicionalmente os conglomerados identificados pelo modelo Poisson com restrição de 10% da população envolveram muitos bairros cujo risco no conglomerado foi menor do que os riscos observados nos conglomerados identificados com outros valores para a restrição. Desta forma, o modelo Poisson com 7% de restrição da população apresentou melhores resultados.

### 3.2.2 Modelo Permutação Espaço-Tempo

A estatística Scan usando o modelo Permutação Espaço-Tempo não identificou conglomerados de homicídios por arma branca no período considerado para nenhum dos percentuais de população utilizados (de 3% a 10%). Para o modelo com restrição de 3% da população, o menor p-valor associado a um conglomerado foi igual a 0,138; para a restrição de 5% o menor p-valor foi 0,267; para 7% foi 0,358 e para 10%, igual a 0,439. Isto indica que segundo o modelo, não há áreas de alto risco ao longo do tempo em qualquer bairro de João Pessoa. Isto pode ser explicado pelo baixo número de óbitos observados nos bairros associado à similaridade no número de ocorrências por bairro. Isto pode levar a não significância estatística em testes que consideram apenas o número de casos, como ocorre com o modelo Permutação Espaço-Tempo. Por outro lado, quando se considera o risco, as informações são ponderadas pelo número de indivíduos expostos. Assim, bairros com mesmo número de casos podem apresentar riscos significativamente distintos quando considerada as respectivas densidades populacionais.

## 4. DISCUSSÕES E CONCLUSÃO

Foi avaliado o desempenho da estatística Scan utilizando os modelos Permutação Espaço-Tempo e Poisson para os homicídios por arma branca ocorridos no município de João Pessoa, PB, nos anos de 2001 a 2009. A comparação foi realizada segundo a abordagem epidemiológica, ou seja, por meio da comparação dos resultados com os mapas de risco.

Entre 2001 e 2005, foram contabilizadas ocorrências em 10% dos bairros no decorrer dos anos, havendo mudanças na configuração espacial dos casos de homicídios no município entre 2001 e 2003. A partir de 2004, foi verificada uma persistência de casos em pelo menos 50% dos bairros em que ocorreram homicídios no ano anterior. De 2006 a 2009 a média anual de bairros com registro de homicídios foi de 27,7%, o que indica que houve maior abrangência na distribuição espacial dos homicídios nos bairros. Adicionalmente, de modo geral, o risco de homicídio por bairro foi, menor do que o observado em anos anteriores.

A comparação dos resultados da estatística Scan espaço-temporal, foi realizada sob a ótica epidemiológica, tendo em vista que não se sabia *a priori* quais e quantos conglomerados havia. Sem a utilização de restrições populacionais, os modelos podem identificar conglomerados correspondentes à totalidade da área de estudo, o

que é inviável na diferenciação de áreas de risco. Os resultados foram comparados com os mapas de risco relativo, sendo esperado que os conglomerados identificados não fossem muito distintos das áreas em que foram registradas ocorrências.

A estatística Scan utilizando o modelo Poisson identificou conglomerados espaço-temporais para todos os percentuais de população. Para as ocorrências registradas entre 2001 e 2005, o modelo com restrição de 3% da população identificou um conglomerado formado apenas pelo bairro com risco mais elevado e seu adjacente. A diferença entre o modelo Poisson com restrição de 3% da população e o modelo com restrição de 5%, foi o acréscimo de três bairros no conglomerado identificado anteriormente, sendo que dois deles não apresentaram registro de ocorrências. Os modelos com 7% e 10% de restrição populacional apresentaram o mesmo resultado: o conglomerado identificado era formado por oito bairros e, embora três deles tenham apresentado risco zero, nos demais foram registrados riscos elevados.

Quanto aos homicídios ocorridos entre os anos de 2006 a 2009, o modelo Poisson com restrição de 3%, 5% e 7% apresentaram o mesmo conglomerado: três bairros foram identificados, sendo que dois dele apresentaram risco elevado e no outro não foram observadas ocorrências. O modelo com 10% da população envolveu sete bairros no conglomerado, sendo que dois deles não foram registrados homicídios por arma branca.

O modelo Permutação Espaço-Tempo não detectou conglomerados espaciais para qualquer percentual de população. A diferença nos resultados é devido a não inclusão da população nos cálculos do modelo Permutação Espaço-Tempo, o que interfere no risco, pois o modelo desconsidera os indivíduos expostos. Assim, quando dois bairros têm densidades populacionais distintas, um modelo que considera apenas o número de casos pode indicar que não há diferença entre eles quando o número de ocorrências é semelhante. Entretanto, modelos que consideram a densidade populacional dos bairros detectam a diferença no risco de ocorrência de eventos em localidades com mesmo número de ocorrências, mas com diferença no número de indivíduos expostos. Desta forma, o modelo Permutação Espaço-Tempo apresenta melhor desempenho em identificar conglomerados quando a distribuição da população exposta ao risco é homogênea.

Como o modelo Permutação Espaço-Tempo considera no cálculo conglomerados que abrangem no máximo 50% do tempo, este mesmo percentual foi utilizado para o modelo Poisson. Esta medida foi tomada a fim de tornar comparáveis os resultados dos mesmos. Entretanto, se utilizada uma restrição de tempo maior, por exemplo, 90%, os mesmos conglomerados tenderão a aparecer, consoante a permanência de riscos elevados nas áreas identificadas com percentuais de população menores.

Com base nos resultados obtidos, o modelo Poisson com restrição populacional de 3% e 5% detectaram poucos bairros com risco elevado entre 2001 e 2005, enquanto os modelos com restrição de 7% e 10% da população obtiveram resultados semelhantes e compreenderam maior número de bairros com alto risco.

Para os anos de 2006 a 2009, o modelo com restrição populacional de 10% selecionou alguns bairros com risco zero, enquanto os modelos com restrição de 3%, 5% e 7%, identificaram apenas três bairros. Adicionalmente, o risco observado no conglomerado obtido do modelo com 10% da população foi igual a 6,24, enquanto no modelo com 7% foi 10,63, que é um risco elevado quando se considera apenas três bairros. Desta forma, o modelo que, ao serem ponderadas a abrangência dos conglomerados identificados e o risco associado, apresentou melhores resultados foi o modelo com restrição de 7% da população.

### REFERÊNCIAS BIBLIOGRÁFICAS

- ASSUNÇÃO, R. M. *Estatística especial com aplicações em epidemiologia, economia e sociologia*. São Carlos, SP: associação Brasileira de Estatística, 2001.
- ASSUNÇÃO, R. M. DUCZMAL, A. R. a simulated annealing strategy for the detection of arbitrary shaped spatial clusters. *Computational Statistics and data analysis*, v. 45, p. 269-286, 2004.
- BALIEIRO, A. S. S. *Detecção de conglomerados dos alertas de desmatamentos no Estado do Amazonas usando estatística de varredura espaço-temporal*. Dissertação (Mestrado em Estatística Aplicada e Biometria) – programa de Pós-Graduação em Estatística Aplicada e Biometria, Universidade Federal de Viçosa, Viçosa, 2008.
- BESAG, J.; NEWELL, J. The detection of clusters in rare diseases. *Journal of the Royal Statistical Society*, v. 154, p. 143-155, 1991.
- COULSTON, J. W. RITTERS, K. H. Geographic analysis of forest health indicators using spatial scan statistics. *Environmental Management*, v. 31, n. 6, p. 764-773, 2003.
- CUZICK, J; EDWARDS, R. Spatial clustering for inhomogeneous populations. *Journal of the Royal Statistical Society B*, 52, p. 73-104, 1990.
- GAWRYSZEWSKI, V. P.; KAHN, T.; JORGE, M. H. P. Informações sobre homicídios e sua integração com o setor saúde e segurança pública. *Rev. Saúde Pública*, v. 39, n. 4, p. 627-633, 2005.
- GÓMEZ-RUBIO, V.; FERRÁNDIZ-FERRAGUD, J.; LÓPEZ, A. Detecting clusters of disease with R. *Journal of Geographical Systems*, v. 7, n. 2, p. 189-206, 2005.
- GUIMARÃES, J. M. X. Estudo epidemiológico da violência por arma branca no município de Porto Grande, Amapá. *Ciência & Saúde Coletiva*, v. 10, n. 2, Abr/Jun 2005.
- KULLDORFF, M. A spatial scan statistics. *Communications in Statistics – Theory and Methods*, v. 26, p. 1481-1496, 1997.
- YAMADA, I.; ROGERSON, P. A.; LEE, G. DeoSutveillance: a GIS-based system for the detection and monitoring of spatial clusters. *Journal of Geographical Systems*, v. 11, p. 155-173.

- KULLDORFF, M. et al. A space-time permutation scan statistic for the early detection of disease outbreaks. *PLoS Medicine*, v. 2, p. 216-224, 2005.
- KULLDORFF, M. et al. An elliptic spatial scan statistic. *Statistics in Medicine*, v. 25, p. 3929-3943, 2006.
- KULLDORFF, M. INFORMATION MANAGEMENT SERVICES, Inc. *SaTScan v8.0: Software for the Space-Time Scan Statistics*. [S.l.], February 2009.
- KULLDORFF, M.; NAGARWALLA, N. Spatial disease cluster: detection and inference. *Statistics in Medicine*, v. 14, p. 799-810, 1995.
- LUCENA, S. E. F.; MORAES, R. M. Análise do desempenho dos métodos Scan e Besag e Newell para identificação de conglomerados espaciais do dengue no municípios de João pessoa entre os meses de janeiro de 2004 e dezembro de 2005. *Boletim de Ciências Geodésicas*, v. 15, n. 3, p. 544-561, 2009.
- MEDRONHO, R. A. *Epidemiologia*. 2 ed. [S.l.]: Ateneu, 2004.
- MOURA, F. R. *Detecção de clusters espaciais via algoritmo scan multi-objetivo*. Dissertação (Mestrado em Estatística) – Universidade Federal de Minas Gerais, Belo Horizonte, 2006.
- NEIL, D. B.; MOORE, A. W.; SABHNANI, M; DANIEL, K. Detection of emerging space-time clusters. In: *Proceedings of the eleventh ACM SIGKDD international conference on Knowledge discovery (KDD '05)*, Chicago (USA), August 21-24, p. 218-227, 2005
- PATIL, G. P.; TALLIE, C. Upper level set scan statistic for detecting arbitrary shaped hotspots. *Envir Acol Stat*, v. 11, p. 183-197, 2004.
- PELLEGRINI, D. C. P. *Análise espaço-temporal da leptospirose no município do Rio de Janeiro (1995-1999)*. Dissertação (Mestrado) – Fundação osvaldo Cruz, Rio de Janeiro, 2001.
- PHEBO, L. Brasil: as armas e as vítimas. In: . RJ: *ISER/7 letras*, 2005. Cap. Impacto da arma de fogo na saúde da população no Brasil.
- R DEVELOPMENT CORE TEAM. *R: A language and Environment for Statistical Computing*. Vienna, Austria, 2008. Disponível em: <<http://www.r-project.org>>.
- ROTHMAN, K. J.; GREENLAND, S. (Ed.) *Modern Epidemiology*. [S.l.] Lippincott-Raven, 1998.
- TANGO, T. A test for spatial clustering adjusted for multiple testing. *Statistics in Medicine*, v. 19, p. 191-204, 2000.
- TANGO, T.; TAKAHASHI, K. A flexibly shaped spatial scan statistic for detecting clusters. *International Journal of Health Geographics*, v. 4, n. 11, 2005.
- WAISELFISZ, J. J. *Mapa da violência 2010 – Anatomia dos Homicídios no Brasil*. [S.l.], 2010.

(Recebido em julho de 2012. Aceito em novembro de 2012.)