

Taxas bayesianas para o mapeamento de homicídios nos municípios brasileiros

Bayesian rates for homicide mapping in Brazilian municipalities

Alexandre Xavier Ywata de Carvalho ¹
 Gabriela Drummond Marques da Silva ¹
 Gilberto Rezende de Almeida Júnior ¹
 Pedro Henrique Melo de Albuquerque ¹

Abstract

This study analyzes homicide incidence per municipality (county) in Brazil in the year 2008. The authors estimate and compare homicide rates according to different methods, finding evidence that depending on the method employed, the results can differ significantly, especially for small municipalities. Bayesian spatial procedures were employed, allowing minimization of variation in the rate estimates. The methods consider a priori distributions and information on contiguity of municipalities. According to the findings, the impact of corrective procedures is not relevant for large municipalities, but such estimates present significant differences for small municipalities. Comparing the different estimates, the authors conclude that there may be distortions in the rates published in the literature. To overcome such potential distortions, it is necessary to take the main goal in each analysis into account. When the emphasis is on overall visualization of the homicide phenomenon, the best option is spatial corrections. However, to obtain more accurate local estimates, Bayesian methods are more appropriate.

Homicide; Mortality Rate; Statistical Analysis

Introdução

O mapeamento da taxa de ocorrência de eventos é uma importante ferramenta para o direcionamento de políticas públicas. Apesar disso, o tema não tem recebido a devida atenção por parte de alguns pesquisadores. De fato, em muitos trabalhos, os índices utilizados apresentam sérios problemas quando as unidades geográficas possuem uma pequena população sob risco e quando o número de casos observados é muito baixo, conforme será visto mais adiante. Quando os municípios são pequenos, a estimação de taxas brutas é pouco indicada. Por serem pouco populosos, o baixo número de casos observados leva a estimativas pouco representativas ou até distorcidas e, por isso, com alta variabilidade ¹.

Esse problema foi solucionado, primeiramente, por meio de técnicas de suavização Bayesiana ². Além dessa proposta, outra possibilidade é a estimativa de taxas mais estáveis, nas quais a vizinhança do município é incluída como informação adicional ³. Tais metodologias são chamadas de empíricas uma vez que a informação *a priori* é construída com base na própria amostra. Pringle ⁴ faz uma revisão de técnicas bayesianas empíricas tendo como base estudos de caso de câncer e mortalidade neonatal.

Algumas áreas já vêm incorporando análises empíricas para aprimorar o mapeamento de taxas em seus estudos. Asiimwe et al. ⁵ utilizam técnicas de mapeamento em pequenas áreas

¹ Instituto de Pesquisa Econômica Aplicada, Brasília, Brasil.

Correspondência
 A. X. Y. Carvalho
 Instituto de Pesquisa Econômica Aplicada,
 SBS, Qd. 1, Bl. J, Ed. BNDES,
 sala 717, Brasília, DF
 70076-900, Brasil.
 alexandre.ywata@ipea.gov.br

para direcionar intervenções na redução de mortalidade de crianças com idade abaixo de cinco anos em Uganda. Já Souza et al.⁶ utilizaram-se dos métodos bayesianos empíricos na análise da distribuição espacial de taxas de hanseníase, com o objetivo de identificar áreas de possível sub-registro de casos, enquanto Cavalini & Leon⁷ fizeram uso da mesma técnica para correção de sub-registros de óbitos por causas mal definidas pelo Sistema de Informações Hospitalares do SUS (SIH-SUS).

Para o cálculo de taxas mapeadas no espaço, diversos *softwares* estão atualmente disponíveis. O *software* estatístico R (The R Foundation for Statistical Computing, Viena, Áustria; <http://www.r-project.org>) possui uma grande variedade de funções já programadas, o que permite ao usuário o cálculo de diferentes estimadores para taxas de algum evento estudado. Rotinas para o cálculo dos estimadores descritos neste artigo podem ser encontradas também no *software* GEODA (Spatial Analysis Laboratory, University of Illinois, Urbana-Champaign, Estados Unidos). O Instituto de Pesquisa Econômica Aplicada (IPEA) desenvolve o *software* IpeaGEO (<http://www.ipea.gov.br/ipeageo>), uma ferramenta de análises estatísticas, com foco em análises espaciais. Na versão 2.0, a ser lançada em 2012, serão incluídas rotinas para o cálculo de taxas bayesianas.

O objetivo deste trabalho é apresentar uma aplicação dos métodos de estimação de taxas para calcular taxas de homicídios nos municípios brasileiros em 2008. Na seção *Métodos* serão apresentados os estimadores bayesianos e espaciais de taxas. A aplicação dos estimadores discutidos neste artigo para o problema de estimação das taxas de homicídios municipais é apresentada na seção *Resultados*. Finalmente, a última seção apresenta a discussão do artigo.

Métodos

Esta seção descreve os principais problemas associados ao cálculo de taxas no formato habitual. Apresenta, em seguida, alguns modos alternativos de estimação.

Taxa bruta

A forma mais usual de mapeamento de taxas utiliza a chamada taxa bruta, dada pela expressão:

$$Tb_i = \frac{Y_i}{P_i} \times 100.000, \quad i = 1, \dots, n \quad (1)$$

sendo i o número de casos observados, i o tamanho total da população sob risco e n o número de unidades geográficas (por exemplo, municípios)

contidas no mapa. Por intermédio desse cálculo, obtém-se o número observado de eventos no município i ($i = 1, \dots, n$) em uma população base de 100 mil pessoas (esse padrão é utilizado, por exemplo, para a apresentação de taxas de homicídios, conforme será apresentado no estudo de caso mais adiante, neste texto).

Para as situações em que o evento de interesse é raro, como número de homicídios (em países desenvolvidos) ou ocorrência de câncer, o mapeamento deve ser mais cuidadoso. Nesses casos, muitas vezes a taxa bruta pode ser igual a zero embora seja razoável dizer que a probabilidade de que a taxa real seja exatamente igual a zero é muito baixa. Tal valor igual a zero pode ser resultado simplesmente de não termos uma janela temporal longa o suficiente, por exemplo. Outro problema marcante ocorre quando a população sob risco, i , é muito pequena e, por isso, a variabilidade das estimativas se torna muito alta⁸. Por exemplo, considere duas regiões: uma com mil habitantes e apenas uma morte e outra com 10 milhões de habitantes e mil mortes; a primeira região pode ser considerada, devido puramente ao acaso, mais violenta do que a segunda se usarmos a taxa bruta.

Taxa espacial

Tendo em conta que o fator espacial possui influência sobre o evento de interesse, esse componente pode desempenhar um papel importante na correção das estimativas. De acordo com essa suposição, espera-se que os valores dos riscos das unidades geográficas correspondam a uma superfície que varia suavemente ao longo do território. Para obter uma estimativa suavizada, a taxa espacial, substitui-se o valor de cada localidade pela taxa encontrada para a região em que ela está localizada. Tal região é composta pelo próprio município e seus vizinhos⁹.

Taxa bayesiana empírica e bayesiana empírica espacial

Existem também soluções próprias da estatística bayesiana que são propícias para enfrentar os problemas mencionados para a taxa bruta. A taxa de mortalidade dos vizinhos (contíguos ou não) do município i é relevante e pode ser utilizada como informação *a priori*. Essa taxa é denominada empírica, pois se utiliza os próprios dados para construir *a priori*.

A taxa bayesiana empírica³ de um município é uma soma ponderada entre a sua taxa bruta e a taxa média global de todo o mapa. Tal soma possui fator de ponderação inversamente proporcional à população em risco P_i ; ou seja, quanto

maior a população de um município, mais confiável é a sua estimativa via TB_i e por isso a taxa bayesiana empírica será aproximadamente igual a TB_i . Entretanto, quando a população é pequena, o valor estimado será aproximadamente igual à taxa global. Tal medida é também chamada taxa bayesiana empírica global.

A taxa bayesiana empírica diminui a variabilidade das estimativas ao restringir a flutuação aleatória ocasionada pelos eventos raros. Essa taxa, contudo, pode ser ainda mais aprimorada com a inclusão de efeitos espaciais de vizinhança em seus cálculos. Para que isso seja feito, um recurso possível é convergir os valores encontrados para uma média local, que será a média dos vizinhos. Logo, a estimativa apresentará uma suavidade espacial condizente com a realidade de muitos eventos. Assim, tem-se a taxa bayesiana empírica espacial, ou taxa bayesiana empírica local, que mantém a mesma estrutura de soma ponderada, substituindo-se, no entanto, a média global pela média local.

Taxa bayesiana de Clayton & Kaldor

A taxa empírica proposta por Clayton & Kaldor ² também utiliza a estatística bayesiana na sua composição. Ela, porém, se destaca das demais porque estabelece uma estrutura clara para a distribuição do risco relativo θ_i , utilizada como informação *a priori*. Algumas distribuições possíveis são os modelos *log-normal* e *gama*, para os quais os cálculos podem ser vistos no artigo original. No segundo modelo, assume-se que as variáveis aleatórias θ_i são independentes e seguem uma distribuição *gama* com parâmetro de escala α e parâmetro de forma v ; ou seja, com média v/α e variância v/α^2 . Considerando a distribuição já mencionada de Y_i condicionada a θ_i , pode-se mostrar que a distribuição *a posteriori* é

também *gama* com parâmetro de escala ($E_i+\alpha$) e parâmetro de forma (Y_i+v). Portanto, a esperança *a posteriori* é:

$$E(\theta_i | Y_i, \alpha, v) = \frac{Y_i + v}{E_i + \alpha} \tag{2}$$

Note que o valor encontrado irá variar entre o SMR (Y_i/E_i) e $\hat{\nu}/\hat{\alpha}$, a média estimada da distribuição de θ_i . A esperança *a posteriori* de θ_i dado Y_i é a estimativa empírica bayesiana final.

Resultados

Nesta seção será analisada a distribuição dos homicídios nos 5.566 municípios brasileiros, no ano de 2008. Todos os dados estão disponíveis na base de dados do Sistema de Informações sobre Mortalidade do Departamento de Informática do SUS (SIM-DATASUS). Nesse sistema, optou-se por utilizar as siglas da Classificação Internacional de Doenças (CID) cujos óbitos foram motivados por causas que sugerem homicídios ou que foram comprovadamente ocasionadas por homicídios. Nesse grupo, enquadram-se as mortes por agressão, negligência e abandono, disparo de arma de fogo e por contato com objeto cortante penetrante. Os códigos relacionados são X85 a Y09 e Y22 a Y24 e Y28 e Y29 da CID-10, como utilizados em Castro et al. ¹⁰. A Tabela 1 é apresentada com as medidas descritivas das taxas calculadas neste artigo, utilizando-se o *software* IpeaGEO.

Se comparado com outros países, pode-se concluir, por meio da Tabela 1, que a média simples das taxas brutas de homicídios no Brasil é elevada. Nos Estados Unidos, por exemplo, o número de mortes por 100 mil habitantes foi de seis pessoas em 2006. Embora alta, a taxa bruta do Brasil, pode estar sendo subestimada já que, em todas as demais taxas, as médias encontradas fo-

Tabela 1

Estatísticas descritivas das taxas de homicídio por município no Brasil em 2008.

	Média	Mediana	Desvio-padrão	Mínimo	Máximo	Coefficiente de variação
Taxa bruta	15,70	10,70	18,74	0,00	211,10	1,19
Taxa bayesiana empírica	21,47	18,80	11,54	2,50	125,90	0,54
Taxa espacial	18,93	14,85	14,11	0,00	109,60	0,75
Taxa bayesiana empírica espacial	17,10	13,70	16,37	0,00	146,80	0,96
Taxa bayesiana Clayton & Kaldor	17,88	14,30	11,42	2,10	120,50	0,64

Nota: taxas por 100 mil habitantes.

Fonte: elaboração dos autores com dados do Departamento de Informática do SUS (DATASUS).

ram superiores a 17 homicídios por 100 mil habitantes. Esse fenômeno pode ter ocorrido em razão do grande número de municípios pequenos no país. Algumas dessas regiões não registraram homicídios, deslocando a média da taxa bruta para valores mais baixos.

Os problemas de estimação da taxa bruta não se restringem à subestimação. Em regiões com um número pequeno de homicídios outro fenômeno verificado é a superestimação. Como pode ser vista, a taxa bruta é a única que qualificou um dos municípios do país como uma região de extremo risco de homicídios, tendo uma taxa de 211,10. O município dito como o mais perigoso do país pela a estimação obtida com a taxa bruta registrou, todavia, apenas cinco homicídios. Esse é o distrito de Nova Marilândia, localizado em Mato Grosso cuja população é de apenas 2.369 pessoas. Segundo a taxa Clayton & Kaldor, entretanto, a estimativa dessa região é reduzida para 57,8 mortes por 100 mil habitantes, enquanto, de acordo com a taxa espacial, esse número é de 35,7.

Logo, os dois problemas mencionados, o de super e subestimação podem ser resumidos em único desafio, a baixa precisão da taxa bruta. A taxa bruta é a única estimativa em que o coeficiente de variação é maior do que 1,0. Como o coeficiente de variação é o desvio padrão dividido pela média, esta é a única taxa em que a variabilidade é maior até do que a estimativa pontual.

Nas análises anteriores, teve-se uma forte evidência de que o tamanho da população está relacionado ao valor das taxas. Nesse caso, é interessante averiguar as particularidades de cada uma das taxas nos maiores e menores municípios do país. A Tabela 2 apresenta os indicadores utilizados neste trabalho, calculados para os vinte menores municípios do Brasil e para as 27 capitais.

Na Tabela 2, é possível observar que a maior parte dos pequenos municípios possui uma taxa bruta igual a zero, mas seria errado inferir que são regiões completamente isentas de risco de homicídios. Como as populações são menores, a estimação do verdadeiro valor é comprometida. Além disso, o sub-registro nas informações do SIM-DATASUS é outro fator que deve ser levado em conta ¹¹. Nesse trabalho, o autor destacou que nos municípios localizados em regiões de difícil acesso aos serviços públicos existem casos de sepultamentos feitos sem a identificação detalhada das causas do óbito.

Dentre os indicadores corrigidos, apenas a taxa bayesiana empírica espacial manteve taxas iguais a zero para alguns municípios. Isso ocorre porque a suavização por esse método é feita pela utilização de uma média local. Caso um municí-

pio esteja localizado em uma região pouco populosa, sem ocorrência de homicídios, a média local pode ser igual a zero. Nessas situações, caso seja do interesse do pesquisador, é possível utilizar vizinhanças de ordem maior do que um, para as quais a suavização será mais acentuada.

A menor redução observada com relação à taxa bruta ocorreu para a taxa bayesiana empírica. Isso pode ser explicado, uma vez que tal taxa possui um método de suavização mais conservador. Como se observa, a taxa bayesiana empírica fez que os indicadores dos menores municípios convergissem para um valor próximo à taxa bruta do Brasil, que é de 26,6 homicídios por 100 mil habitantes. Todos os resultados, no entanto, foram menores do que 26,6, pois a taxa bayesiana empírica corresponde a uma média ponderada entre a taxa bruta do país e a taxa bruta do próprio município.

Na Tabela 2, também são apresentadas as taxas das capitais brasileiras, organizadas de forma ascendente, de acordo com a taxa bruta. Por meio dela, destaca-se, primeiramente, que as três capitais brasileiras mais perigosas, em termos de números de homicídios, fazem parte da Região Nordeste. Esses foram os únicos municípios que registraram taxas superiores a 60 homicídios por 100 mil habitantes. Mesmo para as taxas corrigidas, é possível verificar tal realidade. Comparando-se com o mapa, essas regiões estão localizadas no extremo nordeste do país, na faixa mais escura do mapa que permeia os estados de Pernambuco e Maceió.

As demais regiões que possuem representantes dentre as cinco capitais com maior taxa de homicídio são Norte e Sudeste. Dentre as capitais do Norte, a única que possui um número de mortes maior que 50 por 100 mil habitantes é Belém (Pará). Já no Sudeste, a capital com maior grau de periculosidade é Vitória (Espírito Santo), que apresenta uma taxa de homicídios de quase 60. Em Vitória, a situação é ainda mais alarmante, visto que há uma forte evidência de que as cidades no seu entorno possuem um grau de violência alto. Constata-se o fato por meio da taxa espacial igual a 80,70, muito superior aos demais indicadores observados.

Dentre as capitais com menores níveis de mortalidade por homicídios, destacam-se São Paulo e Palmas (Tocantins). Esses foram os únicos municípios, em meio aos listados, que apresentaram taxas de homicídios inferiores a 20 mortes por 100 mil habitantes. É interessante notar que as taxas corrigidas são concordantes com essa realidade. Nesse caso, São Paulo, que durante a década de 80 foi considerada uma região com nível de violência crescente e preocupante, tem conseguido reverter o cenário ¹².

Tabela 2

Indicadores dos vinte menores municípios e das 27 capitais brasileiras em 2008.

Município	População	Homicídios	Taxa bruta	Taxa bayesiana empírica	Taxa espacial	Taxa bayesiana empírica espacial	Taxa bayesiana Clayton & Kaldor
Borá	834	0	0,0	24,1	10,6	0,0	16,4
Serra da Saudade	889	0	0,0	24,0	12,8	0,0	16,3
Ananguera	1.007	0	0,0	23,6	18,3	0,0	16,1
Nova Castilho	1.111	0	0,0	23,4	0,0	0,0	15,9
Oliveira de Fátima	1.114	0	0,0	23,4	16,0	0,0	15,9
Araguainha	1.138	0	0,0	23,3	21,2	0,0	15,9
Miguel Leão	1.228	0	0,0	23,1	2,5	2,1	15,8
Cedro do Abaeté	1.228	0	0,0	23,1	17,5	0,0	15,8
André da Rocha	1.262	1	79,2	34,0	9,0	15,8	25,4
Chapada de Areia	1.274	0	0,0	23,0	14,4	0,0	15,7
Parari	1.283	0	0,0	22,9	9,2	7,8	15,7
Crixás do Tocantins	1.299	0	0,0	22,9	25,6	0,0	15,6
Lagoa Santa	1.308	0	0,0	22,9	17,0	0,0	15,6
Uru	1.397	0	0,0	22,6	7,4	0,0	15,5
Serra Nova Dourada	1.407	0	0,0	22,6	20,3	14,0	15,5
Nova A. do Ivaí	1.426	0	0,0	22,6	16,6	0,0	15,5
Santa Salete	1.441	0	0,0	22,5	3,8	3,4	15,4
Cachoeira de Goiás	1.441	0	0,0	22,5	23,7	0,0	15,4
Tupirama	1.448	0	0,0	22,5	18,1	0,0	15,4
São Félix do Tocantins	1.451	0	0,0	22,5	4,7	0,0	15,4
Maceió	924.143	933	101,0	100,3	93,9	100,6	100,14
Salvador	2.948.733	1.878	63,7	63,6	65,2	63,7	63,55
Recife	1.549.980	976	63,0	62,8	62,3	62,9	62,70
Vitória	317.817	186	58,5	57,8	80,7	59,8	57,39
Belém	1.424.124	742	52,1	52,0	51,8	52,1	51,88
João Pessoa	693.082	347	50,1	49,8	49,5	49,5	49,65
Curitiba	1.828.092	774	42,3	42,3	48,0	42,4	42,22
Natal	798.065	324	40,6	40,5	39,6	39,6	40,34
Cuiabá	544.737	220	40,4	40,2	42,4	40,8	40,02
Porto Alegre	1.430.220	570	39,9	39,8	43,0	40,0	39,71
Macapá	359.020	143	39,8	39,6	34,6	38,9	39,29
Porto Velho	379.186	147	38,8	38,5	36,9	38,7	38,28
Manaus	1.709.010	648	37,9	37,9	35,4	37,8	37,81
São Luís	986.826	360	36,5	36,4	33,8	36,2	36,31
Goiânia	1.265.394	458	36,2	36,1	35,2	36,1	36,06
Belo Horizonte	2.434.642	879	36,1	36,1	38,0	36,2	36,03
Fortaleza	2.473.614	874	35,3	35,3	35,6	35,4	35,27
Brasília	2.557.158	814	31,8	31,8	38,1	31,9	31,78
Rio de Janeiro	6.161.047	1.821	29,6	29,6	34,2	29,6	29,54
Aracaju	536.785	143	26,6	26,6	31,6	27,1	26,49
Boa Vista	260.930	69	26,4	26,5	26,0	26,0	26,15
Rio Branco	301.398	78	25,9	25,9	22,6	22,6	25,64
Campo Grande	747.189	183	24,5	24,5	24,8	24,8	24,41
Teresina	793.915	178	22,4	22,5	19,7	22,2	22,37
Florianópolis	402.346	89	22,1	22,2	20,1	20,5	22,03
São Paulo	10.990.249	1.701	15,5	15,5	18,7	15,5	15,48
Palmas	184.010	27	14,7	15,2	16,7	16,5	14,82

Nota: taxas por 100 mil habitantes.

Fonte: elaboração dos autores com dados do Departamento de Informática do SUS (DATASUS).

Embora em termos absolutos apresente um dos maiores registros de homicídios listados, percentualmente a capital se caracteriza como uma das mais seguras do país.

Em relação ao grau de suavização das taxas corrigidas, pode-se observar que a maior parte delas não possui mudanças substanciais com relação à taxa bruta. A diferença, nesse caso, não chega a ultrapassar uma unidade, na maior parte das situações. Uma vez que todas as capitais possuem um número alto de habitantes, apenas a taxa espacial é destoante para algumas delas. Essa é a situação, por exemplo, de Maceió (Alagoas), Vitória, Curitiba (Paraná) e Brasília, nas quais a alteração repercutiu em uma variação de pelo menos 6 homicídios por 100 mil habitantes. Assim, a ordenação dos municípios não apresenta modificações bruscas, se arranjada de acordo com qualquer um dos indicadores utilizados.

É possível verificar que as taxas espaciais e bayesianas são importantes para a melhora das

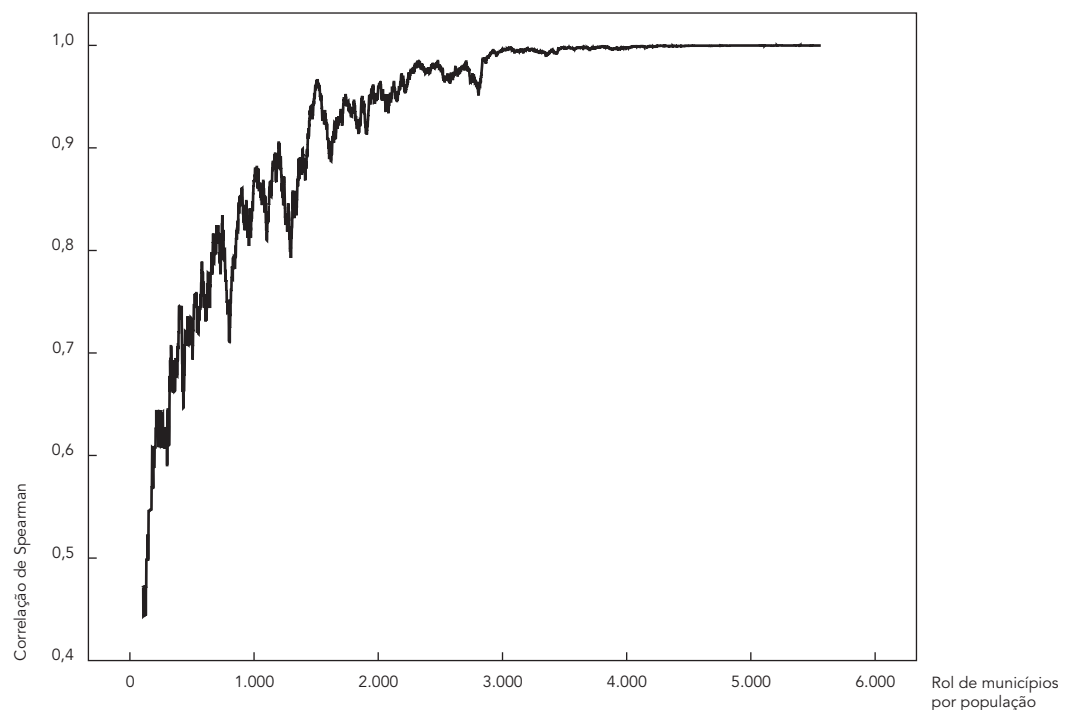
taxas. Essa melhora, entretanto, possui intensidade de suavização direcionada pelo tamanho da população. É interessante então quantificar a relação entre o tamanho do município e a quantidade de correção da taxa bruta. Para isso, foi realizado um ensaio apresentado na Figura 1. Cada gráfico representa a correlação de Spearman entre cada uma das taxas corrigidas e a taxa bruta, em uma janela de 100 municípios.

Segundo a Figura 1, a atuação das correções na taxa bruta em municípios de população pequena é maior do que nos grandes municípios. Verifica-se o padrão crescente nos quatro gráficos. Embora a tendência de crescimento seja encontrada em todos os quatro casos, existem dois comportamentos diferentes quanto à intensidade. A relação entre a taxa bruta e as que possuem um componente espacial apresenta um crescimento aproximadamente linear. A correlação entre a taxa bruta e as taxas não espaciais, por sua vez, possuem um ritmo de crescimento

Figura 1

Correlação de Spearman entre a taxa bruta e as demais taxas para uma janela de 100 municípios.

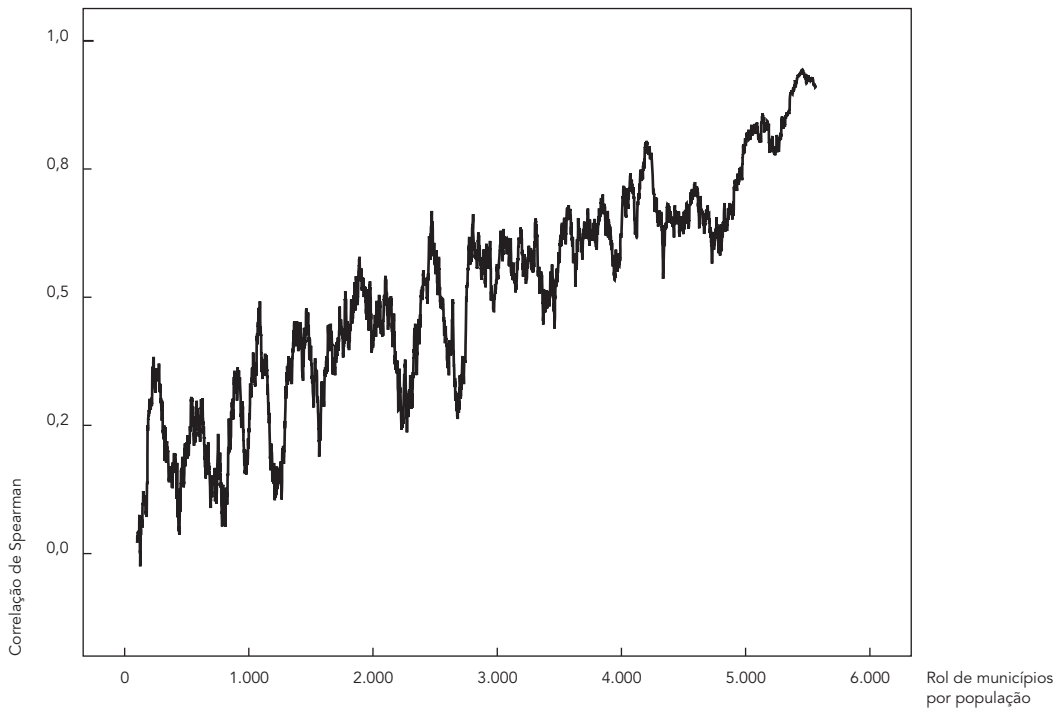
1a) Correlação de taxa bruta com taxa bayesiana empírica



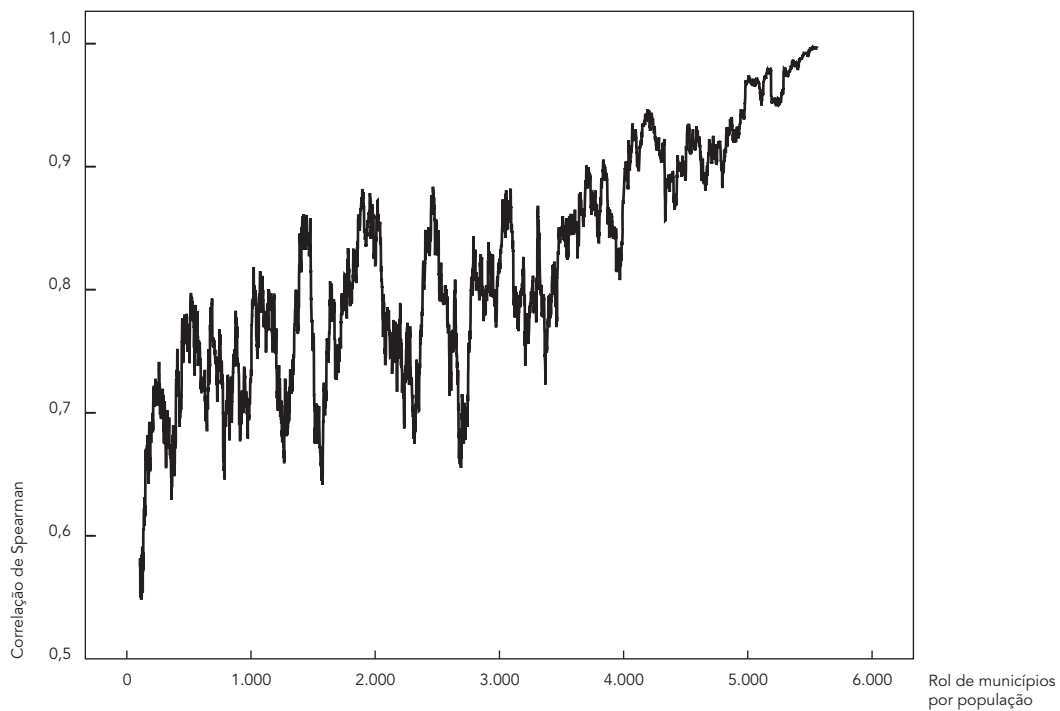
(continua)

Figura 1 (continuação)

1b) Correlação de taxa bruta com taxa espacial



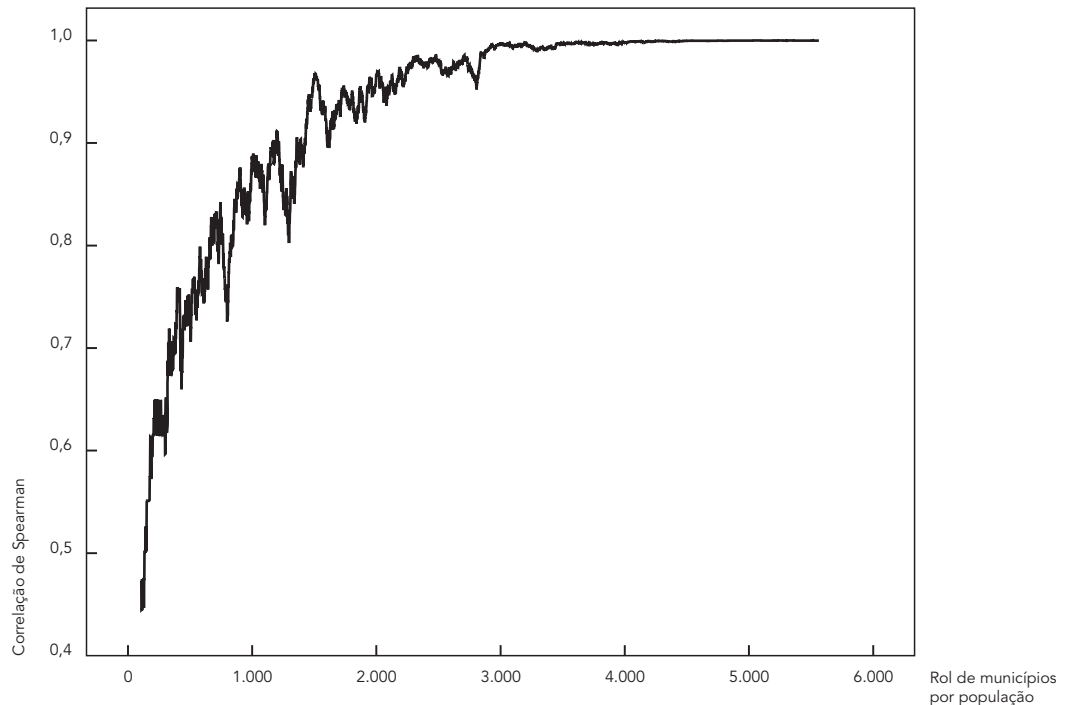
1c) Correlação de taxa bruta com taxa bayesiana empírica espacial



(continua)

Figura 1 (continuação)

1d) Correlação de taxa bruta com taxa bayesiana Clayton & Kaldor



mais rápido, por conta de um comportamento aparentemente *log-linear*, o que significa que todas as taxas fazem correções maiores para os municípios pequenos. Os indicadores não espaciais, contudo, levam à menor suavização das taxas para cidades de porte médio do que os indicadores espaciais.

Quanto à intensidade da associação, pode-se dizer que quase todas as taxas possuem uma correlação alta com a taxa bruta, independente do tamanho da população. Porém, isso não é verdade para a taxa espacial, cuja associação é pequena para tamanhos de amostras pequenos. Como a taxa espacial é uma média da taxa do município e de seus vizinhos, a taxa bruta do próprio município tende a possuir pouca influência no resultado final. Para as demais taxas, a correlação alta com a taxa bruta é uma constatação positiva, pois indica que elas corrigem algumas distorções da taxa bruta, sem distorcê-la completamente. Já a taxa espacial parece ser uma boa estimativa apenas globalmente, pois faz que as taxas apresentem apenas variações suaves no espaço. Localmente, a taxa espacial pode apresen-

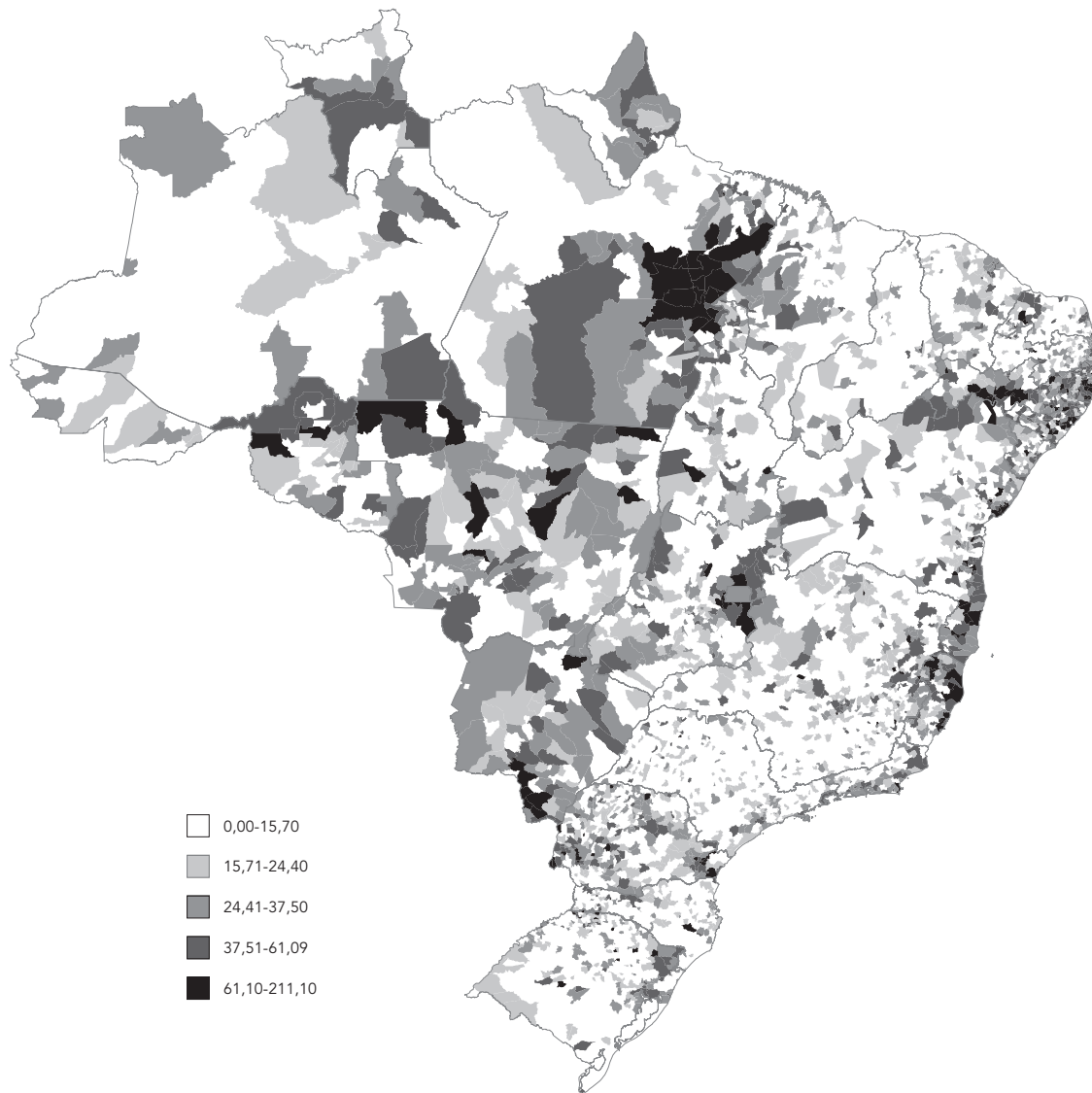
tar resultados não condizentes com a realidade dos pequenos municípios.

Uma vez que o tamanho do município é componente do peso utilizado na taxa bayesiana empírica e na taxa bayesiana de Clayton & Kaldor, a suavização, nesse caso, é função do tamanho do município. Com isso, segundo a Figura 1, é possível observar que a variabilidade em torno da média é menor para as taxas não espaciais. Por conseguinte, municípios de mesmo tamanho podem apresentar taxas com graus diferentes de suavização de acordo especialmente com as taxas espaciais. Assim, caso o pesquisador opte pela influência do espaço, o tamanho da população terá uma participação minimizada na correção da taxa. Para uma visualização do fenômeno discutido, as Figuras 2 e 3 apresentam a distribuição de cada uma das taxas para os municípios brasileiros em 2008.

Na Figura 2, importa destacar a presença de uma grande quantidade de valores extremos, representados pela categoria mais escura, se comparado com os mapas da Figura 3. Por mais que também fosse esperada uma presença maior dos

Figura 2

Taxa bruta de homicídios nos municípios brasileiros em 2008.



valores na classe inferior, isso não ocorre. Provavelmente, os valores extremos inferiores permaneceram na classe mais clara, mesmo após correção, já que as suavizações não foram muito bruscas em sua maioria. Neste mesmo mapa pode-se verificar que, no Norte, a presença de grandes municípios leva a leituras erradas da realidade na região. É o caso do Município de Altamira (Pará), que é o maior município do país, e como está na quarta classe, pode induzir a uma

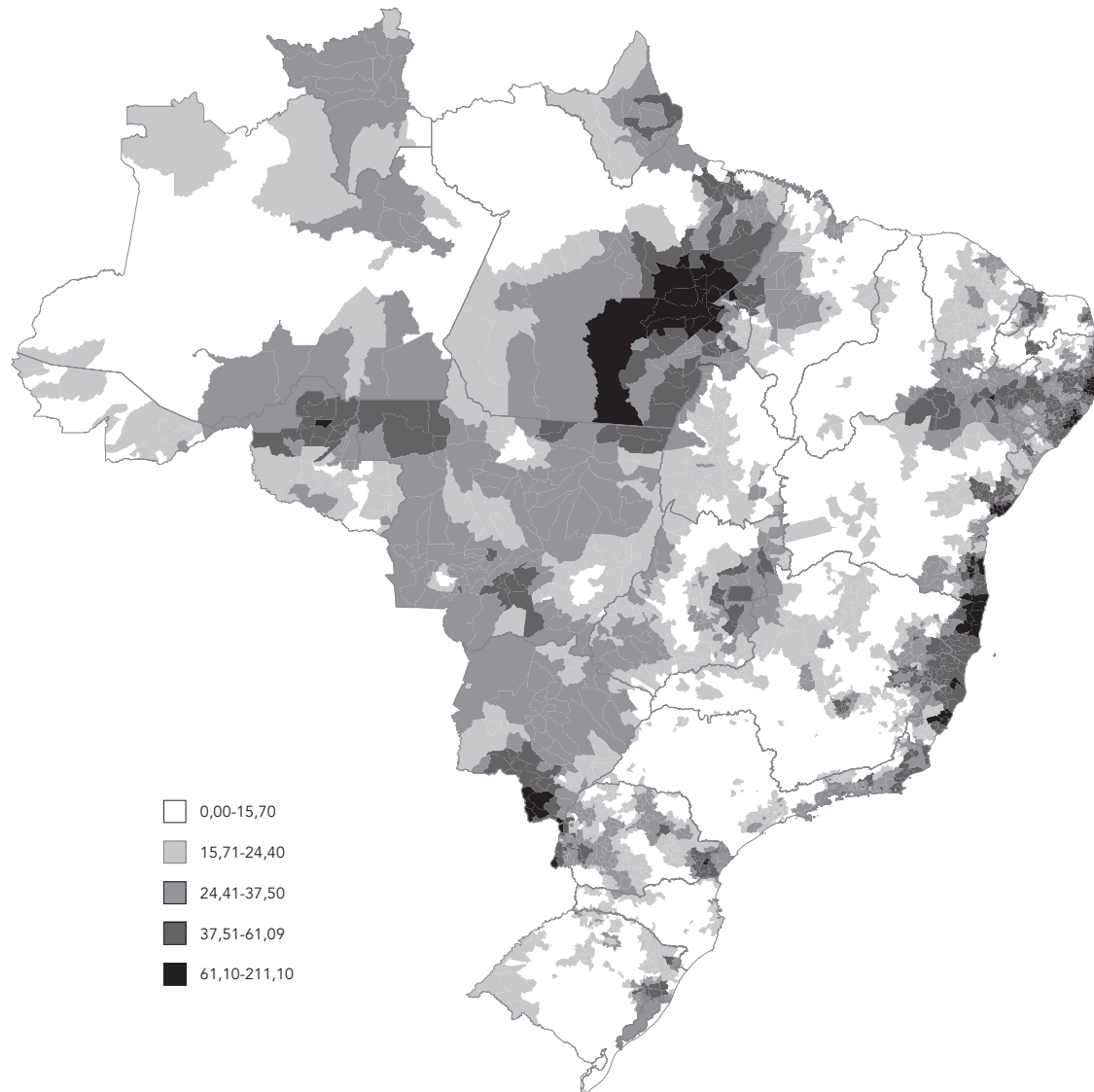
falsa impressão de maior periculosidade no estado como um todo.

Na Figura 3, é possível observar, primeiramente, que o mapa da taxa espacial é o que possui maior diferença em relação aos demais padrões. Isso se deve ao fato de que ele utiliza somente o fator espacial para a correção das taxas. Dessa forma, a visualização dos dados é priorizada porque as taxas se modificam suavemente no espaço. Como atualmente o Brasil é formado

Figura 3

Taxas espacial, bayesiana empírica, bayesiana empírica espacial e bayesiana empírica Clayton & Kaldor de homicídios nos municípios brasileiros em 2008.

3a) Taxa espacial



(continua)

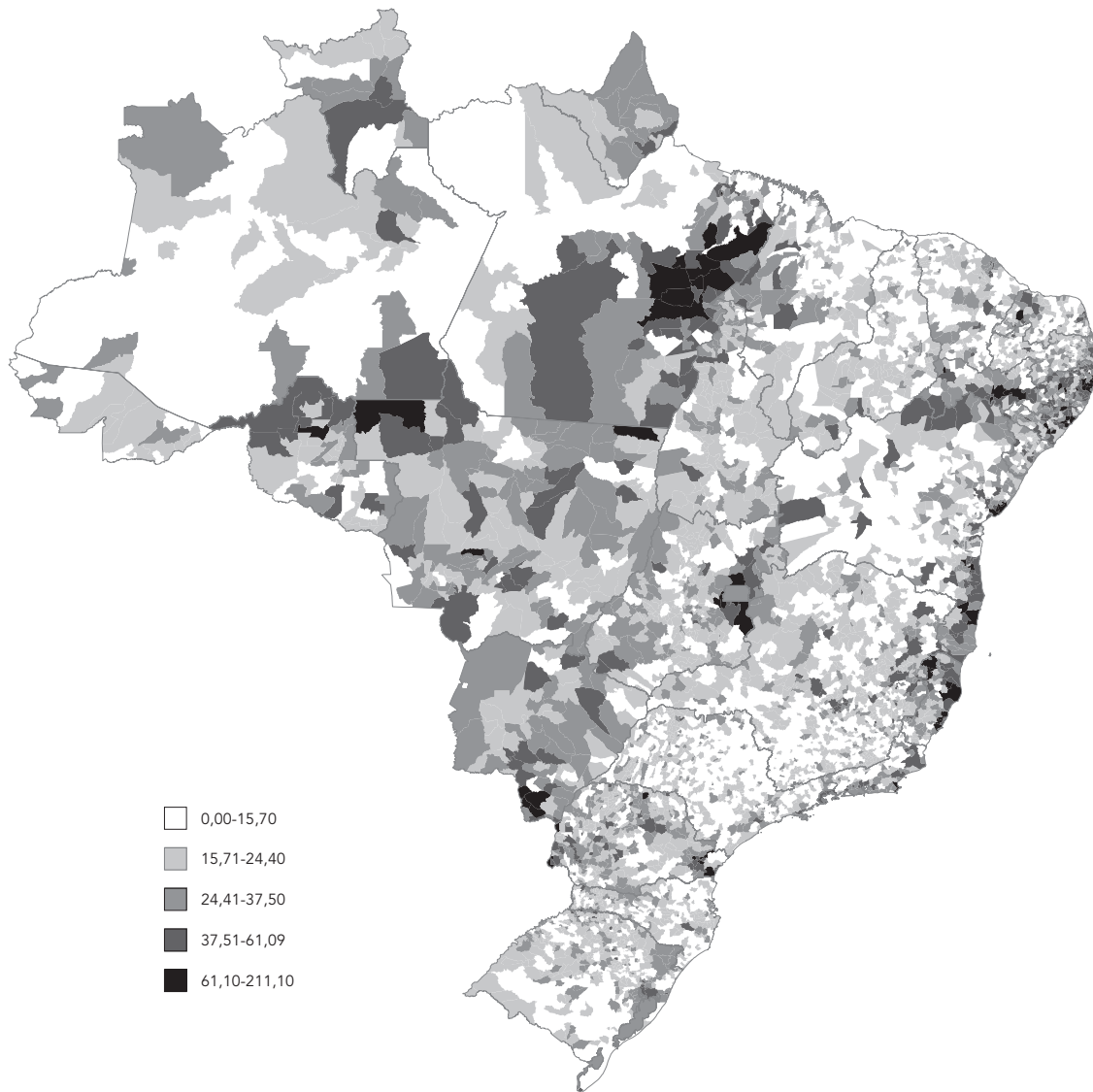
por 5.566 pequenos municípios, a variação é ainda mais branda. Além disso, no litoral norte do Espírito Santo e sul da Bahia, os municípios de Guarapari e Linhares (Espírito Santo) e Porto Seguro (Bahia) apresentam grande influência na taxa espacial, aumentando a taxa dos municípios vizinhos.

O mapa da taxa empírica bayesiana apresenta muitos municípios na classe que varia de

24,41 a 37,50, o que indica a influência da média geral que é 26,70 homicídios por 100 mil pessoas. Nota-se ainda que os mapas baseados na taxa bayesiana empírica e na taxa bayesiana empírica espacial são muito semelhantes. O mapa com o componente espacial, no entanto, parece reproduzir o mapa da taxa bayesiana empírica de forma suavizada. Em todas as situações, percebe-se que a taxa bayesiana empírica e a taxa bayesiana

Figura 3 (continuação)

3b) Taxa bayesiana empírica



(continua)

de Clayton & Kaldor são muito parecidas. Isso indica que a média global é próxima da esperança *a priori*.

Discussão

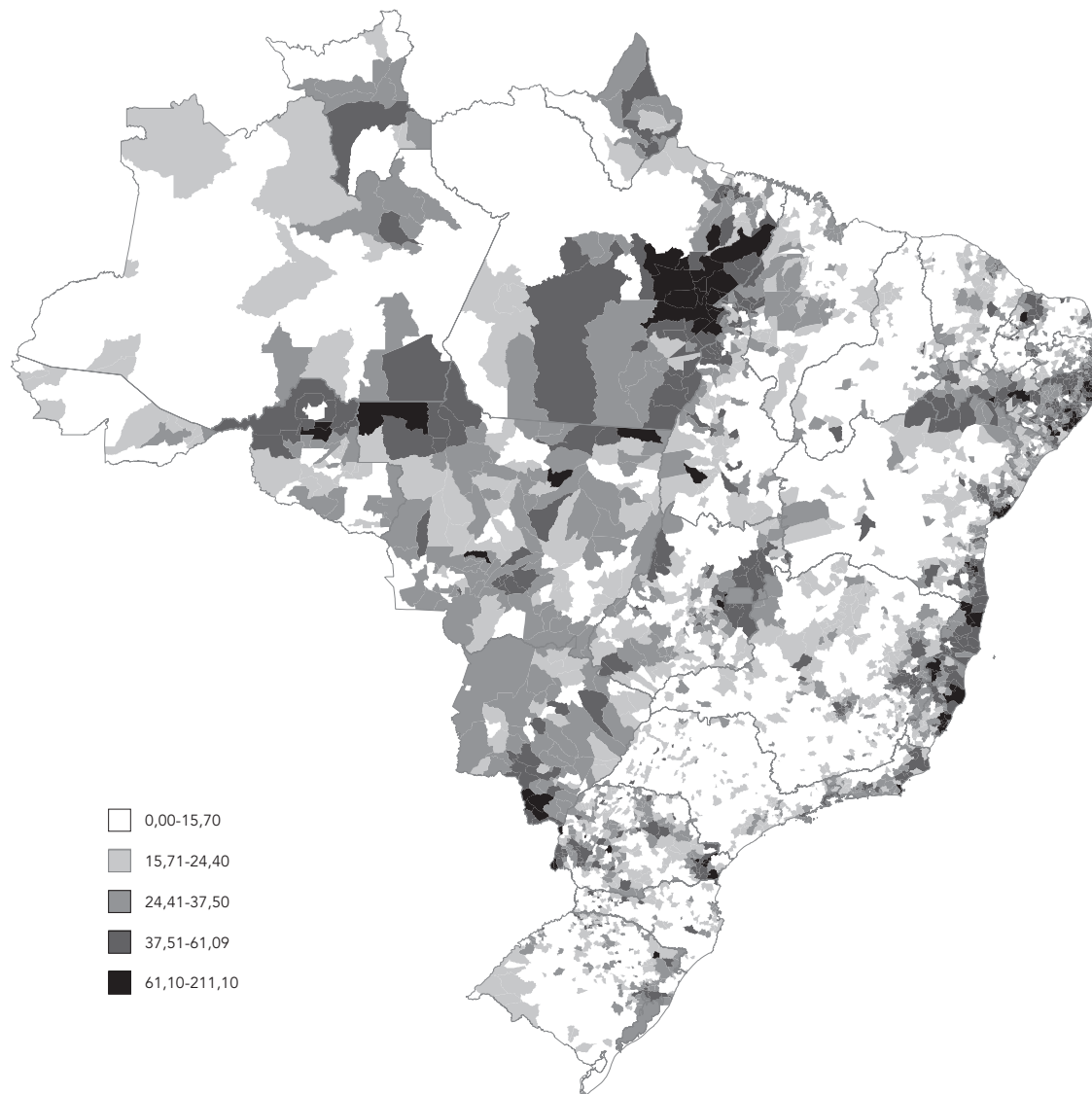
Neste trabalho, foi feita uma revisão de algumas das principais técnicas de aprimoramento de dados de taxas. Mostrou-se a importância dos trabalhos de mapeamento, que são úteis para

direcionamento de políticas públicas relativas à segurança e saúde. Frente à relevância da área a utilização de taxas brutas, foi considerada pouco indicada por causa da sua alta variabilidade aleatória, particularmente quando parte das unidades geográficas estudadas possui populações muito baixas. Os métodos sugeridos, em contraposição, suavizaram os valores extremos por meio de técnicas bayesianas e espaciais.

Diante disso, um exemplo da funcionalidade das medidas empíricas foi feito com dados de

Figura 3 (continuação)

3c) Taxa bayesiana empírica espacial

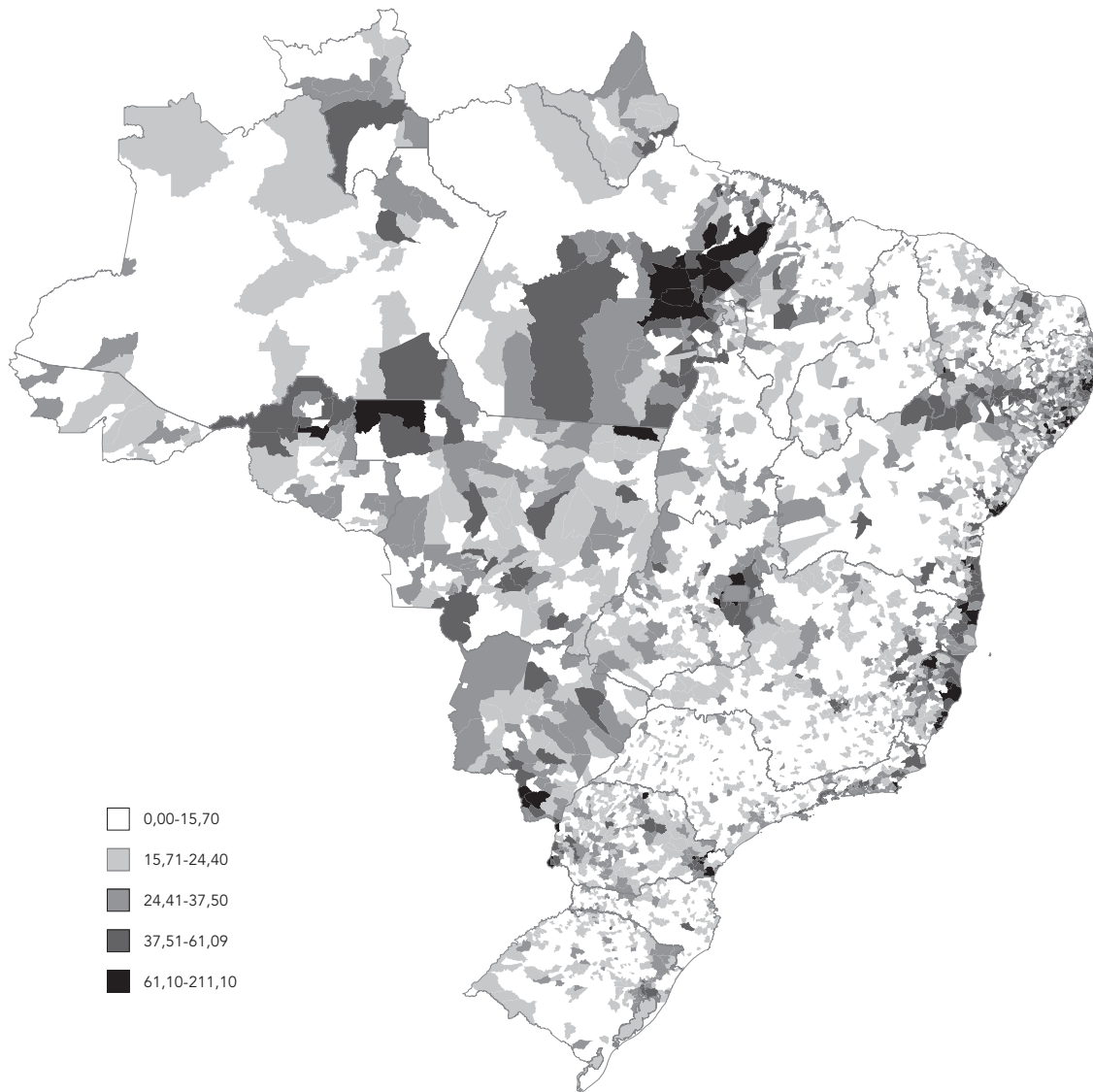


homicídio no Brasil em 2008. Em um país com extensão continental como o Brasil, onde as diferenças regionais são bruscas, não se aconselha a aplicação da taxa bayesiana empírica, pois a correção dela é feita com base na média do país. Melhor opção é a taxa bayesiana empírica espacial que atribui maior influência aos municípios geograficamente próximos, apresentando resultados mais coerentes regionalmente.

A taxa espacial é apenas a média dos vizinhos do município, incluindo o município, e é utilizada para obter uma superfície suavizada do evento, sendo por isso pouco útil para obter valores pontuais. No caso desta aplicação, as taxas que melhor se adaptaram foram as bayesianas, empírica, empírica espacial e Clayton & Kaldor. No caso da taxa bayesiana de Clayton & Kaldor, a distribuição *a priori* influencia muito as estima-

Figura 3 (continuação)

3d) Taxa bayesiana de Clayton & Kaldor



tivas para os municípios pequenos, e temos que verificar se a distribuição *a priori* utilizada faz sentido na aplicação.

Portanto, a escolha do indicador a ser utilizado dependerá do intuito a ser atingido com o trabalho. Caso o mapa utilizado apresente muitos polígonos e grandes diferenças regionais, como é o caso do Brasil, aconselha-se a utilização da taxa bayesiana empírica espacial ou a taxa de Clayton

& Kaldor. A taxa espacial é mais indicada se o objetivo for a análise visual dos dados. Se, entretanto, deseja-se obter estimativas mais apuradas, desconsiderando-se o efeito do espaço, deve-se utilizar a taxa de Clayton & Kaldor, que possui menor variabilidade.

Resumo

Este trabalho analisa a ocorrência de homicídios por município, em todo o Brasil, no ano de 2008. O texto apresenta uma comparação entre as taxas de mortalidade obtidas por diferentes metodologias e levanta evidências de que, de acordo com o método utilizado, os resultados diferem entre si, especialmente para municípios pequenos. Foram utilizados métodos de estatística Bayesiana e espacial como possibilidade para minimização da variabilidade entre as taxas obtidas. Isso foi feito por meio da incorporação de distribuições a priori ou de informações sobre vizinhança dos municípios no cálculo das taxas de mortalidade. Observou-se que o impacto dos procedimentos corretivos não foi relevante para municípios grandes. Por outro lado, tais medidas apresentaram grandes mudanças nas taxas dos pequenos municípios. Os valores calculados indicam possíveis distorções das taxas divulgadas usualmente para o país. Para corrigi-las, é necessário ter em vista o objetivo da análise. Caso o enfoque do trabalho seja a visualização global de um fenômeno, deve-se optar pelas correções espaciais. Se, entretanto, desejar-se obter uma estimativa local mais apurada, o mais indicado é a utilização dos métodos Bayesianos.

Homicídio; Taxa de Mortalidade; Análise Estatística

Colaboradores

A. X. Y. Carvalho foi responsável pela concepção do artigo, coordenou a equipe e colaborou com a análise e revisão do trabalho. P. H. M. Albuquerque participou das tarefas de coordenação da equipe, programação dos algoritmos e análise e revisão do artigo. G. D. M. Silva e G. R. Almeida Júnior contribuíram com a coleta dos dados, programação dos algoritmos e trabalharam com a redação do texto e análise dos dados.

Agradecimentos

Os autores agradecem aos comentários de Maurício Mota Saboya Pinheiro, Marcelo Abi-Hamia Caetano e Daniel Ricardo de Castro Cerqueira. Os erros remanescentes são de responsabilidade dos autores apenas. Este texto não necessariamente reflete o posicionamento do Instituto de Pesquisa Econômica Aplicada (IPEA), sendo de total responsabilidade dos autores.

Referências

1. Catelan D, Lagazio C, Biggeri A. A hierarchical Bayesian approach to multiple testing in disease mapping. *Biom J* 2010; 52:784-97.
2. Clayton D, Kaldor J. Empirical bayes estimates of age-standardized relative risks for use in disease mapping. *Biometrics* 1987; 43:671-81.
3. Marshall RJ. Mapping disease and mortality rates using empirical Bayes estimators. *J R Stat Soc Ser C Appl Stat* 1991; 40:283-94.
4. Pringle DG. Mapping disease risk estimates based on small numbers: an assessment of empirical Bayes techniques. *Econ Soc Rev* 1996; 27:341-63.
5. Asiimwe JB, Jehopio P, Atuhaire L, Mbonye A. Examining small area estimation techniques for public health intervention: lessons from application to under-5 mortality data in Uganda. *J Public Health Policy* 2011; 32:1-14.
6. Souza WV, Barcellos CC, Brito AM, Carvalho MS, Cruz OG, Albuquerque MFM, et al. Aplicação de modelo Bayesiano empírico na análise espacial da ocorrência de hanseníase. *Rev Saúde Pública* 2001; 35:474-80.
7. Cavalini LT, Leon ACMP. Correção de sub-registros de óbitos e proporção de internações por causas mal definidas. *Rev Saúde Pública* 2007; 41:85-93.
8. Rao JNK. *Small area estimation*. New Jersey: Wiley; 2003.
9. Kafadar K. Smoothing geographical data, particularly rates of disease. *Stat Med* 1996; 15:2539-60.
10. Castro MSM, Assunção RM, Durante MO. Comparação de dados sobre homicídios entre dois sistemas de informação, Minas Gerais. *Rev Saúde Pública* 2003; 37:168-76.
11. Waiselfisz JJ. *Mapa da violência 2010: anatomia dos homicídios no Brasil*. São Paulo: Instituto Sangari; 2010.
12. Peres MFT, Vicentin D, Nery MB, Lima RS, Souza ER, Cerda M, et al. Queda dos homicídios em São Paulo, Brasil: uma análise descritiva. *Rev Panam Salud Pública* 2011; 29:17-26.

Recebido em 01/Ago/2011

Versão final reapresentada em 16/Dez/2011

Aprovado em 19/Mar/2012