

Distribuição espacial do risco: modelagem da mortalidade infantil em Porto Alegre, Rio Grande do Sul, Brasil

Spatial risk distribution: modeling infant mortality in Porto Alegre, Rio Grande do Sul State, Brazil

Silvia E. Shimakura ¹
Marília Sá Carvalho ²
Denise R. G. C. Aerts ^{3,4}
Rui Flores ⁴

¹ Centro Politécnico, Setor de Ciências Exatas, Departamento de Estatística, Universidade Federal do Paraná, C. P. 19081, Curitiba, PR 81531-990, Brasil.
² Departamento de Epidemiologia e Métodos Quantitativos em Saúde, Escola Nacional de Saúde Pública, Fundação Oswaldo Cruz, Rua Leopoldo Bulhões 1480, Rio de Janeiro, RJ 21041-210, Brasil.

³ Programa de Pós-Graduação em Saúde Coletiva, Universidade Luterana do Brasil, Rua Miguel Toste 101, Canoas, RS 92420-280, Brasil.

⁴ Secretaria Municipal de Saúde de Porto Alegre, Av. João Pessoa 325, Porto Alegre, RS 90040-000, Brasil.

Abstract Estimation and mapping of risk profiles are the main concerns of epidemiology. This paper analyzes spatial distribution of infant mortality cases as compared to live-born controls from Porto Alegre, Rio Grande do Sul. The modeling framework adopted in this research work is a spatial point process. Under this structure, a risk measure which continuously varies over the study region is defined and estimated using generalized additive model methods. This approach has the advantage of allowing for risk factors that are simple and easy to interpret. The procedure also allows the construction of tolerance contours which help identify areas of significantly high/low risk and an overall test for the null hypothesis of constant risk over the region. Application of this method to infant mortality data showed a highly significant spatial variation in risk for neonatal mortality data and non-significant results for post-neonatal mortality data.

Key words Spatial Analysis; Epidemiologic Models; Infant Mortality

Resumo Estimaco e mapeamento de perfis de risco so interesses da Epidemiologia. Neste trabalho,  analisada a distribuico espacial de casos de mortalidade infantil, comparados a controles de nascidos vivos amostrados do Sistema de Informaces sobre Nascidos Vivos da cidade de Porto Alegre, Rio Grande do Sul, Brasil. A modelagem adotada neste trabalho baseia-se em um processo pontual espacial, na qual se define uma medida de risco que varia continuamente sobre a regio de estudo e estimada por meio de mtodos de modelos aditivos generalizados. Essa abordagem possui a vantagem de permitir a incorporao, no modelo, de efeitos de determinantes individuais e ecolgicos de risco sob forma simples e de fcil interpretao. Tambm permite a construico de contornos de tolerncia que auxiliam na identificao de reas de alto/baixo risco e de um teste global da hiptese nula de risco constante relativa  regio. A aplicao do mtodo aos dados de mortalidade infantil mostrou variao espacial no risco altamente significativa para mortalidade neonatal e no significativa para mortalidade ps-neonatal.

Palavras-chave Anlise Espacial; Modelos Epidemiolgicos; Mortalidade Infantil

Introdução

O efeito das desigualdades sobre a qualidade de vida e, conseqüentemente, nas condições de saúde da população tem sido historicamente um dos objetos centrais da Epidemiologia. No momento atual, a reestruturação da economia internacional – com a globalização, a flexibilização de mercados (Martins & Loureiro, 1996) e seus conseqüentes ajustes – determina paradoxalmente o êxito econômico e o fracasso do bem-estar de grandes segmentos da população. Os efeitos dessa crise mostraram-se intensificados nas áreas metropolitanas e, sobretudo, nas capitais, com queda dos níveis de emprego, concentração de renda e agravamento das desigualdades sociais. Nesse contexto, a epidemiologia vem incorporando, em seu arsenal tecnológico, recursos metodológicos que buscam explicitar a desigualdade da distribuição da saúde e da doença nas sociedades, assim como, principalmente, subsidiar o planejamento das ações do setor saúde, sejam estas de caráter preventivo ou assistencial, voltadas ao ambiente ou aos indivíduos.

Um exemplo da ação desses determinantes estruturais é o efeito do processo da segregação urbana sobre a saúde dos indivíduos. Esse fenômeno expressa-se na expulsão de grupos populacionais para a periferia da cidade carente de infra-estrutura urbana essencial, na exclusão dos benefícios sociais e nos diferentes perfis de morbi-mortalidade desses grupos. O impacto de semelhante processo sobre a saúde dos grupos populacionais pode ser estudado de diferentes formas, desde o indivíduo ao ecológico, sendo particularmente interessante a reunião destes dois olhares e a identificação dos locais em que é mais necessária a intervenção.

Pesquisas realizadas em Porto Alegre, no Rio Grande do Sul, Brasil, estudaram a relação entre segregação urbana e mortalidade infantil (Barcellos et al., 1986; Guimarães & Fishmann, 1986; Oliveira & Simões, 1986; SMSSS, 1989), desnutrição (Aerts, 1992) e retardo no crescimento (Aerts, 1996), encontrando associações significativas entre esses. Estudos como o de Guimarães & Fischmann (1986), utilizando dados de 1980, mostraram que, nas favelas, o coeficiente de mortalidade infantil chegou a 80,7 por 1.000, enquanto que, nas regiões não faveladas, o coeficiente mais alto identificado foi de 45,3.

Os dois componentes da mortalidade infantil – neonatal e pós-neonatal – têm importância variada conforme a situação epidemiológica da área em análise. Maiores proporções de mortalidade no período pós-neonatal rela-

cionadas a causas como doenças respiratórias e diarreias – evitáveis por medidas simples –, ocorrem em situações de grande carência social e precário acesso a serviços de saúde. Ao contrário, quando os níveis de mortalidade infantil atingem valores baixos, predominam causas neonatais, tais como a prematuridade excessiva e as doenças congênitas, mais difíceis de serem evitadas.

Analisando as causas da mortalidade infantil nos anos de 1980 e 1992, Martins & Loureiro (1996) identificaram que as afecções do período perinatal (APPN), seguidas das doenças respiratórias agudas, foram as principais causas de óbito entre os moradores de áreas com infra-estrutura urbana e ocupação regular do solo. Em contrapartida, para os moradores em áreas irregulares, observaram que as doenças respiratórias foram as maiores responsáveis pelos óbitos, superando as APPN. A análise desses achados aponta para uma melhor condição de vida e de saúde da população moradora em áreas regulares em comparação com as áreas irregulares, como seria esperado.

Entretanto, explicitar a interação entre fatores individuais e ambientais, trazendo o “contexto de volta para a Epidemiologia” (Diez-Roux, 1998) exige a aplicação de novas técnicas de análise. Esse é o objetivo central do presente trabalho: utilizar novas técnicas estatísticas para a análise dos dados de mortalidade infantil em Porto Alegre, buscando identificar a forma como se distribuem esses óbitos e seus determinantes de risco no território da cidade.

O trabalho encontra-se dividido em quatro seções. A Seção 2 descreve os dados utilizados neste trabalho e a metodologia estatística empregada. Em função da excepcional qualidade dos dados do Sistema de Informação sobre Nascidos Vivos (SINASC) e do Sistema sobre Mortalidade (SIM), em Porto Alegre, e da existência de um Sistema de Informação Geográfica implantado que permite localizar pontualmente no território do município cada evento de interesse, foi possível empregar uma metodologia estatística proposta recentemente por Kelsall & Diggle (1998) para análise dos dados. Essa metodologia é baseada na teoria de processos pontuais espaciais. Sob esta estrutura de modelagem define-se uma medida de risco que varia continuamente sobre a região e que pode ser estimada por intermédio de métodos de modelos aditivos generalizados.

A definição da medida de risco e a descrição do procedimento de estimação também são explicitados na Seção 2. A aplicação da metodologia aos dados de mortalidade infantil de Porto Alegre é apresentada na Seção 3. Final-

mente, na Seção 4 são discutidos os resultados da análise e feitas considerações gerais acerca do método utilizado.

Metodologia

Os dados

No presente trabalho, foram analisados dados oriundos do SIM e do SINASC, *linkados* na rotina da Equipe de Informação em Saúde, da SMS, incluindo as seguintes variáveis: localização da residência, peso ao nascer, semanas gestacionais e sexo de crianças que nasceram durante o ano de 1998, no Município de Porto Alegre, e que morreram antes de completarem o primeiro ano de vida. Informações a respeito da idade e grau de instrução da mãe, bem como do tipo de gravidez e de parto, também foram incluídos. Do total de 374 óbitos registrados no período, foram retirados 75, por terem nascido no ano anterior, em local de residência fora da região continental de Porto Alegre, pela ausência de informações de covariáveis ou impossibilidade de georreferenciamento, sendo analisados, separadamente, 183 ocorridos no período neonatal (menos de 28 dias) e 116 pós-neonatal (entre 28 dias e 12 meses incompletos), distribuindo-se igualmente as perdas nos dois grupos.

Os denominadores utilizados para o cálculo do risco foram *controles* selecionados aleatoriamente no SINASC, entre as crianças que também nasceram em 1998, considerados representantes da distribuição espacial da população em risco. Foi adotado um esquema de amostragem (2:1), ou seja, dois nascimentos para cada óbito, perfazendo um total de 600 indivíduos.

O pacote estatístico S-Plus, versão 5.5 (Mathsoft Incorporation, 1998) e programas implementados na linguagem Fortran77 foram utilizados nesse trabalho.

Método

O mapeamento de medidas de risco em epidemiologia é usualmente feito a partir de dados agregados por áreas, em geral, divisões administrativas. Embora os mapas resultantes sejam facilmente interpretáveis, com a agregação perde-se a estrutura espacial detalhada dos dados e, conseqüentemente, as variações espaciais em pequena escala não são detectáveis.

Outro problema na análise de dados de contagem é que os modelos usualmente empregados para dados desse tipo – tais como a distri-

buição binomial negativa ou a Poisson mista – carecem de interpretação espacial adequada, já que são baseados em suposição pouco realista de independência entre as contagens nos recortes da região. Portanto, é preferível derivar um modelo estatístico para os dados de contagens a partir da suposição de um processo pontual espacial latente, ao invés de modelar as contagens diretamente (Diggle, 1992). A dificuldade na interpretação surge porque os recortes são invariavelmente definidos em termos de fronteiras políticas arbitrárias em lugar de um conjunto natural de unidades espaciais. Uma mudança nos recortes da região pode introduzir alteração substancial nos mapas resultantes.

A estrutura de modelagem adotada nesse trabalho baseia-se em um processo pontual espacial (Diggle, 1983), pelo qual se pode definir e estimar uma medida de risco que varia continuamente na região de interesse.

• Definição da medida de risco

Para a definição da medida de risco assume-se um desenho caso-controle espacial. No presente estudo, os controles são uma amostra da população em risco: os nascidos vivos no ano de 1998 no Município de Porto Alegre.

Considerando $(s_1, s_2, \dots, s_{n_1})$ as localizações na região de estudo A dos n_1 casos e $(s_{n_1+1}, s_{n_1+2}, \dots, s_n)$ as localizações dos $(n - n_1)$ controles como observações de dois processos Poisson I e II, com intensidades $\lambda_1(s)$ e $\lambda_2(s)$, respectivamente. Define-se o logaritmo do risco na localidade s como

$$\rho(s) = \log\{\lambda_1(s)/\lambda_2(s)\}, \quad (1)$$

e o objetivo na análise é então investigar a variação espacial de $\rho(s)$ em A .

• Estimação do risco

A estimação da medida de risco tem aspectos conceitualmente semelhantes à estimação individual do risco. No entanto, ao invés de se avaliar somente as variações nas frequências de fatores individuais de risco entre casos e controles, avalia-se simultaneamente a variação na distribuição espacial de casos quando comparados à distribuição espacial de controles.

Uma série de trabalhos na área de Epidemiologia ambiental que têm como tema a estimação espacial do risco (Bithell, 1992; Kelsall & Diggle, 1995a, 1995b; Lawson & Williams, 1993). Kelsall & Diggle (1998) consideram o uso de regressão binária não-paramétrica e apresentam uma proposta semi-paramétrica na estimação

da razão de densidades. O presente trabalho utiliza essa abordagem, em particular, a semi-paramétrica também conhecida como abordagem de Modelos Aditivos Generalizados (GAM), já que essa tem a vantagem de permitir a estimação espacial do risco controlando por fatores individuais de risco.

Considere-se novamente $(s_1, s_2, \dots, s_{n_1})$ as localizações dos n_1 casos e $(s_{n_1+1}, s_{n_1+2}, \dots, s_n)$ as localizações dos $(n - n_1)$ controles. Seja y_i um indicador associado ao ponto s_i tal que $y_i = 1$ se o indivíduo i for um caso ou 0 caso contrário. Assume-se que y_i ($i = 1, \dots, n$) são realizações de variáveis aleatórias independentes $Y_i \sim \text{Bernoulli}(p(s))$, onde:

$$p(s) = \frac{q_1 \lambda_1(s)}{q_1 \lambda_1(s) + q_2 \lambda_2(s)} \quad (2)$$

Na equação (2) acima, q_1 representa a proporção de casos registrados e q_2 a de controles, em relação ao total de fato existente, incorporando ao modelo a possibilidade de sub-registro, desde que a perda seja aleatória no espaço. Sob esta suposição, tem-se que:

$$\log \left\{ \frac{p(s)}{1 - p(s)} \right\} = \rho(s) + c, \quad (3)$$

onde, $c = \log(q_1/q_2)$, ou seja, c é meramente uma constante aditiva e, portanto, não modifica as características da distribuição espacial do risco sobre a região. Assim, pode-se obter estimativas espaciais do risco definido em (1), à parte de uma constante aditiva, a partir de um modelo simples atribuído a variáveis binárias. Como regra prática, recomenda-se a utilização de escala logarítmica na base dois, já que, nessa escala, o aumento de uma unidade na superfície de log-risco de uma localização para outra implica risco dobrado.

A inclusão de efeitos de covariáveis no modelo, como sugerido por Kelsall & Diggle (1998), é dada por:

$$\log \left\{ \frac{p(s, x)}{1 - p(s, x)} \right\} = \beta x + g(s), \quad (4)$$

onde x é o vetor de covariáveis, β são seus efeitos e a única suposição acerca de g é que esta é uma função suave, porém desconhecida, das coordenadas espaciais s . Em particular, se o risco é assumido constante na região, então $g(s) = 0$ e o modelo (4) reduz-se simplesmente a um modelo de regressão logística usual (Hosmer & Lemeshow, 1989). O modelo (4), portanto, nada mais é do que um modelo de regressão logística estendido por uma componente aditiva $g(s)$ que, por suposição, varia suavemente no espaço.

O procedimento de estimação de β e $g(s)$ baseia-se em métodos iterativos usuais de mo-

delos aditivos generalizados (Hastie & Tibshirani, 1990) que pode ser resumido no seguinte algoritmo:

Passo 1: Faz-se $g(s) = 0$ e obtém-se as estimativas iniciais dos coeficientes β ajustando um modelo de regressão logística

$$\log \left\{ \frac{p(x)}{1 - p(x)} \right\} = \beta x,$$

(Hosmer & Lemeshow, 1989).

Passo 2: Calcula-se $\hat{\eta}_i = \hat{\beta}x_i + \hat{g}(s)$, $\hat{p}_i = \exp(\hat{\eta}_i) / [1 + \exp(\hat{\eta}_i)]$ e $z_i = \hat{\eta}_i + (y_i - \hat{p}_i) / \hat{p}_i (1 - \hat{p}_i)$

Passo 3: Ajusta-se um modelo aditivo da forma $Z = \beta x + g(s) + \varepsilon$:

a) Estima-se $g(s)$ usando regressão kernel ponderada (Wand & Jones, 1995) de $u_i = z_i - \hat{\beta}x_i$ em s_i com pesos $w_i = \hat{p}_i (1 - \hat{p}_i)$:

$$\hat{g}(s) = \frac{\sum_{i=1}^n w_i K\left(\frac{s - s_i}{h}\right) \cdot u_i}{\sum_{i=1}^n w_i K\left(\frac{s - s_i}{h}\right)},$$

onde $K(\cdot)$ é a função Kernel definida nesse trabalho pela função de densidade da distribuição normal bivariada padrão. O parâmetro h , doravante denominado banda, influencia o grau de suavidade na superfície de risco estimada – quanto maior o valor de h , maior o alisamento.

b) Estima-se β por métodos de mínimos quadrados ponderados de $z_i - \hat{g}(s_i)$ em u_i com pesos w_i .

c) Repita os passos (i) e (ii) até convergência das estimativas.

Passo 4: Repita os passos 2 e 3 até convergência das estimativas.

• Escolha do valor de banda

O critério para a escolha de h varia conforme a distribuição espacial dos dados e dos objetivos da análise. Essa eleição pode ser guiada por questões de interesse prático, como, por exemplo, ressaltar variações locais do risco entre os bairros de uma cidade. Neste caso, o valor de h deve refletir o diâmetro médio das áreas de interesse. Evidentemente essa escolha também depende do espaçamento dos dados, já que a interpretação da superfície estimada pode ser comprometida em regiões onde a banda escolhida é pequena em relação à distância entre pontos observados.

Alternativamente, podem ser utilizados métodos automáticos para seleção de h baseados em critérios estatísticos (Fan et al., 1995; Härdle, 1990; Kelsall, 1996; Kelsall & Diggle, 1998; Wand & Jones, 1995). Esses métodos são

conhecidos na literatura estatística como *métodos de validação cruzada*, cujo princípio é selecionar o valor de h que minimiza determinada função dos dados e dos valores preditos pelo modelo. O método utilizado no presente trabalho, proposto por Kelsall & Diggle (1998), usa validação cruzada de mínimos ponderados para o passo de regressão não-paramétrica. No passo de suavização 3(i), escolhe-se o valor de h que minimiza

$$CV(h) = \frac{\sum_{i=1}^n w_i \{z_i - \hat{g}^{-1}(s_i)\}^2}{n},$$

onde $\hat{g}^{-1}(s_i)$ é a estimativa de $g(s_i)$ construída com o valor de banda h , usando todos os dados com exceção do par (s_i, z_i) . Estas podem ser comparadas, como foi feito neste trabalho.

• Significância

Além do mapeamento do risco, é importante avaliar se a superfície estimada varia significativamente na região, ou seja, se existem evidências estatísticas suficientes para rejeitar a hipótese nula de risco constante na região, tendo-se controlado por fatores individuais de risco. Em termos do modelo (4), isso equivale ao teste da hipótese $H_0: g(s) = 0$.

Também é de interesse a construção de contornos de tolerância que auxiliam na identificação de áreas onde o risco é significativamente superior (ou inferior) à média global. Ou seja, reconhecendo o papel de dado fator como preditor importante da mortalidade infantil e controlando por este, deseja-se identificar áreas nas quais o risco é significativamente elevado, buscando orientar a intervenção.

O teste global do risco e a identificação de áreas de baixo e alto risco podem ser feitos utilizando o método de simulação Monte Carlo, proposto por Kelsall & Diggle (1998).

a) Geram-se dados consistentes com H_0 , ajustando-se um modelo de regressão logística convencional aos dados e calculando-se as probabilidades ajustadas \hat{p}_i . Fixando-se as localizações de cada ponto, amostra-se n_1 dos n indivíduos sem reposição com probabilidade proporcional a \hat{p}_i e estes são rotulados como casos e os $n-n_1$ restantes como controles.

b) Calcula-se nova estimativa de $g(s)$, $\hat{g}_1(s)$, a estimativa centralizada em torno da média $\bar{g}_1(s) = \hat{g}_1(s) - \bar{g}_1$, onde

$$\bar{g}_1 = n^{-1} \sum_{i=1}^n \hat{g}_1(s_i) \text{ e a estatística } t_1 = n^{-1} \sum_{i=1}^n (\hat{g}_1(s_i))^2.$$

Repetem-se os passos 1 e 2 m vezes.

c) Constrói-se uma superfície de p-valores

que para cada s fornece a proporção dos valores de $\hat{g}_j(s)$, $j=1, \dots, m$, menores do que a estimativa original, diga-se $\hat{g}_0(s)$.

d) Adicionam-se os contornos de 0.05 e 0.95 da superfície de p-valores ao mapa de $\hat{g}_0(s)$ como contornos de 90% de confiança para indicar áreas de alto/baixo risco.

e) Para o teste de hipótese, define-se k o número de $t_j > t_0$ (obtida a partir de $\hat{g}_0(s)$) e o nível de significância correspondente por $p = (k + 1)/(m + 1)$.

Resultados

Para o estudo da variação espacial no risco de morte infantil em Porto Alegre no ano de 1998, foram realizadas análises separadas para cada um dos tipos de mortalidade infantil considerados.

O modelo ajustado assume um efeito linear do total de sete covariáveis no logaritmo do risco. Duas delas são características pessoais da criança (*sexo* e *peso*); duas são características da mãe (*idade* e *instrução*); duas são relacionadas à gravidez (*tipo de gravidez* e *idade gestacional*); e a última refere-se ao tipo de parto (*tipo de parto*). Segue abaixo o modelo ajustado:

$$\log \left\{ \frac{p(s, x)}{1 - p(s, x)} \right\} = \beta_0 + \beta_1 \text{sexo} + \beta_2 \text{peso} + \beta_3 \text{idade} + \beta_4 \text{inst} + \beta_5 \text{ges} + \beta_6 \text{grav} + \beta_7 \text{parto} + g(s).$$

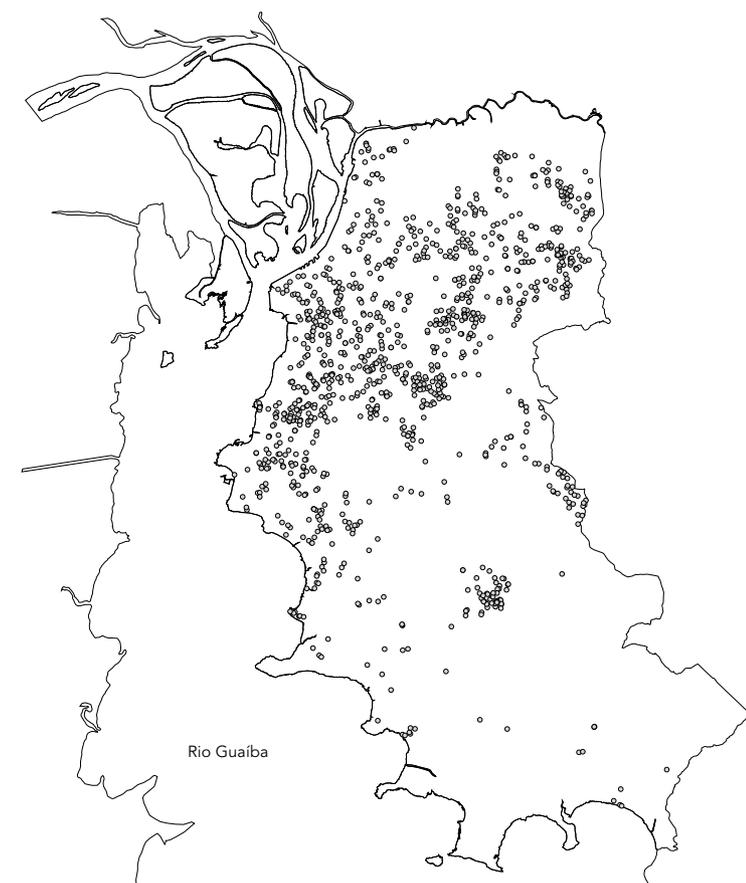
Com exceção das covariáveis *peso ao nascer* e *idade materna no nascimento da criança*, que são contínuas, as demais covariáveis foram categorizadas da seguinte forma: sexo (masculino = 1 e feminino = 0), instrução (1º grau incompleto = 1 e 1º grau completo ou mais = 0), tipo de gravidez (gemelar = 1 e não gemelar = 0), duração da gestação (prematuro = 1 ou a termo = 0) e tipo de parto (operatório = 0 e vaginal = 1).

Para o ajuste do modelo utilizou-se o parâmetro de suavização $h = 2000\text{m}$, visando identificar variações locais espaciais do risco ao nível de bairros cujo diâmetro médio aproximado é de 2km. O teste global de risco constante e a identificação de regiões de risco significativamente (90%) alto ou baixo foram obtidos para h estimado por validação cruzada.

A Figura 1a apresenta a distribuição espacial da população em risco de mortalidade infantil em Porto Alegre, no ano de 1998. Na Figura 1b encontram-se os óbitos neonatais e pós-neonatais, onde cada ponto no mapa representa a residência de um caso. Como se pode observar, o mapa de pontos dos óbitos, apesar do impacto visual imediato, reproduz a distribui-

Figura 1a

Distribuição da amostra de nascidos vivos. Porto Alegre, 1998.



○ nascidos vivos (amostra)

ção da população em risco que, invariavelmente, apresenta ocupação espacial determinada por aspectos geográficos e econômicos.

Note-se a heterogeneidade da distribuição da população em risco na região representada na Figura 1a. Uma fração considerável da população concentra-se nos bairros da região Norte e Noroeste de Porto Alegre, enquanto que o restante encontra-se irregularmente distribuído nos demais bairros. Fica evidente, portanto, a necessidade de uma análise que considere esse padrão heterogêneo de ocupação, o que se obtém utilizando o método descrito anteriormente. As estimativas dos parâmetros do modelo, usando-se um valor de banda $h = 2.000m$ e controlando-se pelo fator espacial, são apresentadas na terceira coluna da Tabela 1.

Dentre as características individuais da criança, mais uma vez se confirma a importância do baixo peso ao nascer na determinação do óbito tanto neonatal como pós-neonatal. A prematuridade mostrou-se associada à mortalidade neonatal, e a instrução da mãe, à pós-neonatal. Marginalmente significativas são as variáveis tipo de parto e tipo de gravidez. A primeira, indicando um efeito protetor do parto vaginal para a neonatal e a segunda, um efeito protetor da gestação não-gemelar para pós-neonatal.

As demais variáveis não apresentaram associação estatisticamente significativa com a mortalidade infantil entre a população estudada.

As estimativas da componente não-paramétrica do modelo $g(s)$, equivalentemente interpretadas como superfícies de risco estimadas, controlando-se por fatores individuais, são apresentadas nas Figuras 2a e 2b.

A análise indica variação espacial significativa no risco de mortalidade neonatal sobre a cidade de Porto Alegre, com p-valor de 0,024, e identifica áreas de alto (áreas dentro das linhas brancas) e uma de baixo (área dentro da linha negras) risco na região. Quanto ao risco de mortalidade pós-neonatal, embora seja elevado em algumas áreas, parece não variar significativamente sobre a região (p-valor = 0,137). Nesse caso, não se identificam regiões de significativo baixo risco e, por isso, não há áreas delimitadas por linhas negras.

Com o objetivo de avaliar o impacto do parâmetro espacial de alisamento nas estimativas dos efeitos dos fatores em estudo e no teste global do risco, o mesmo modelo foi novamente ajustado, mas, dessa vez, com valores de banda obtidos utilizando o método de validação cruzada (descrito na Seção 3) – respectivamente $h_{vc} = 3.000m$ para a mortalidade neonatal e $h_{vc} = 1.857m$ para a pós-neonatal. Os resultados são apresentados na sexta coluna do Tabela 1. De maneira geral, as estimativas dos parâmetros permaneceram inalteradas ao variar o valor de h . Ou seja, entre as variáveis selecionadas, a mortalidade neonatal tem como determinantes individuais o peso ao nascer e a idade gestacional, e a pós-neonatal, além do peso, mostra também a instrução da mãe.

As diferenças entre ambas as análises são mais visíveis nos mapas de risco expostos na Figura 3a e 3b. Como esperado, a superfície de risco estimada para mortalidade neonatal com valor de banda obtido pelo critério de validação cruzada é mais suave do que a anterior, já que o valor de h é maior do que o utilizado anteriormente. Um efeito contrário, porém menos acentuado, é observado na superfície de

risco estimada para a mortalidade pós-neonatal, uma vez que o novo valor de h é próximo ao anterior. A análise estatística dessas superfícies indica a existência de variação espacial ainda mais significativa no risco de mortalidade neonatal (p -valor = 0,012) e revela significância (marginal) no risco de mortalidade pós-neonatal (p -valor = 0,084). Esses resultados são explicados pelo fato de que a precisão associada ao mapa de risco estimado com valor de banda obtido pelo método de validação cruzada é maior (ou no mínimo igual) àquele obtido por critérios subjetivos, o que facilita a detecção de variações espaciais do risco pelo teste estatístico.

Como anteriormente, é possível identificar áreas de alto (áreas dentro das linhas brancas) e de baixo (áreas dentro das linhas negras) risco de mortalidade infantil em determinadas regiões. Note que um número menor de áreas de alto risco são determinadas na segunda análise. Essa diferença é mais marcante no risco de mortalidade neonatal (Figura 3a).

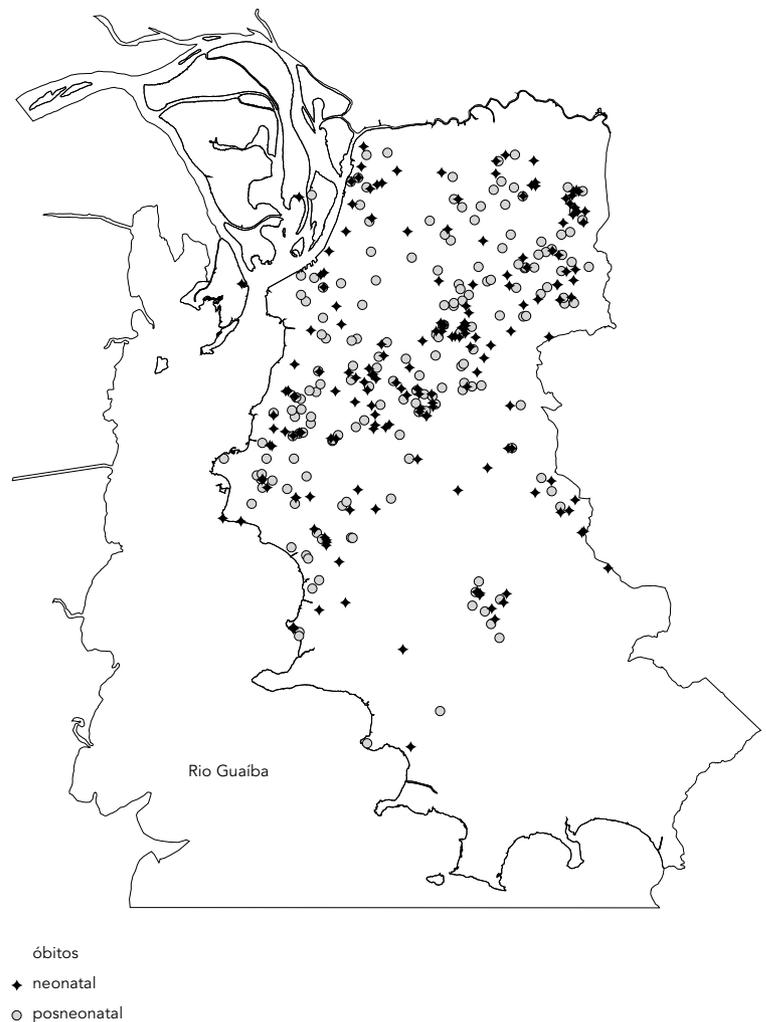
Discussão

Os modelos usuais de análise dos determinantes da mortalidade infantil permitem estimar efeito de covariáveis individuais. Por sua vez, os modelos mais empregados de análise espacial tratam apenas ecologicamente as variáveis, não sendo possível controlar pelas variações individuais. Além disso, a disponibilidade usual dos dados espaciais por áreas administrativas diminui a resolução espacial. Este trabalho, baseado na qualidade e disponibilidade das informações de saúde da Secretaria Municipal de Saúde de Porto Alegre, mostrou ser possível obter estimativas do risco continuamente no espaço, controlando por fatores individuais de risco e utilizando as localizações pontuais de casos e controles.

Para a mortalidade neonatal, os dois principais determinantes individuais foram o peso ao nascer e idade gestacional, como esperado. Entretanto, esses são pouco influenciados pelos diferenciais sócio-econômicos existentes nos atuais patamares de mortalidade infantil das grandes cidades brasileiras (Campos et al., 2000). No caso da mortalidade pós-neonatal, além do já referido efeito do baixo peso ao nascer, evidenciou-se também o efeito da escolaridade materna, ou seja, as chances de sobrevivência das crianças são em grande parte determinadas por sua inserção social, ao passo que a baixa escolaridade da mãe, neste estudo, representou, para a criança, um excesso de risco para a mortalidade no período pós-neonatal.

Figura 1b

Distribuição da localização dos óbitos, neonatais e pós-neonatais.
Porto Alegre, 1998.



No entanto, raramente encontra-se essa característica isolada. Sabe-se que muitas dessas mães iniciaram sua vida reprodutiva precocemente, quando ainda eram adolescentes, estando despreparadas para a difícil tarefa da maternagem, tendo que abandonar os estudos ou o trabalho. Tiveram também muitos filhos, com pequeno intervalo interpartal, o que dificulta, ainda mais, a qualidade da atenção dedicada às crianças. Certamente, a ação conjunta desses fatores também influencia as chances de sobrevivência dessas crianças.

O que esse modelo traz de novo é que, por meio dele, é possível identificar as regiões de

Tabela 1

Estimativas dos efeitos de covariáveis utilizando o valor da banda selecionada ($h = 2.000$) e o da calculada pelo método da validação cruzada ($h = h_{vc}$).

Mortalidade/Fator	$h = 2.000$			$h = h_{vc}$		
	Estimativa	Erro padrão	P-valor	Estimativa	Erro padrão	P-valor
Neonatal						
Idade da mãe	-0,0109	0,0195	0,5742	-0,0131	0,0197	0,5059
Instrução da mãe	0,0546	0,2731	0,8415	0,0718	0,2753	0,7942
Duração da gestação	1,1540	0,3710	0,0019	1,1685	0,3737	0,0018
Tipo de gravidez	-0,1860	0,6517	0,7754	-0,2006	0,6558	0,7598
Tipo de parto	-0,5284	0,2822	0,0616	-0,5320	0,2838	0,0613
Intercepto	4,0680	0,9333	0,0000	4,0717	0,9487	0,0000
Sexo	-0,3711	0,2695	0,1690	-0,3674	0,2713	0,1761
Peso ao nascer	-0,0019	0,0002	0,0000	-0,0018	0,0002	0,0000
Pós-neonatal						
Idade da mãe	-0,0171	0,0171	0,3187	-0,0171	0,0171	0,3174
Instrução da mãe	0,8046	0,2427	0,0010	0,8031	0,2422	0,0010
Duração da gestação	0,5560	0,3663	0,1295	0,5617	0,3654	0,1247
Tipo de gravidez	0,9682	0,5465	0,0769	0,9734	0,5451	0,0745
Tipo de parto	0,0606	0,2529	0,8106	0,0586	0,2523	0,8163
Intercepto	2,0012	0,8776	0,0229	2,0134	0,8747	0,0216
Sexo	0,3027	0,2309	0,1902	0,2996	0,2303	0,1938
Peso ao nascer	-0,0014	0,0002	0,0000	-0,0014	0,0002	0,0000

sobre-risco, onde a mortalidade infantil é significativamente maior, mesmo controlando pelas demais variáveis. Essas superfícies de risco diferenciado relacionam-se às condições sócio-econômicas das áreas, independente da ocorrência individual de baixo peso e prematuridade. Isso é, as famílias moradoras nessas regiões estão expostas a situações de risco coletivo que não aquelas indicadas pelas variáveis do SINASC.

Embora o presente estudo não tenha sido planejado com o objetivo de investigar o complexo universo da determinação da mortalidade, muito menos em sua dimensão individual, os resultados aqui apresentados podem contribuir para futuros estudos a respeito desse tema. É oportuno lembrar que a mortalidade infantil foi utilizada como exemplo de evento que pode ser investigado com o auxílio da análise espacial.

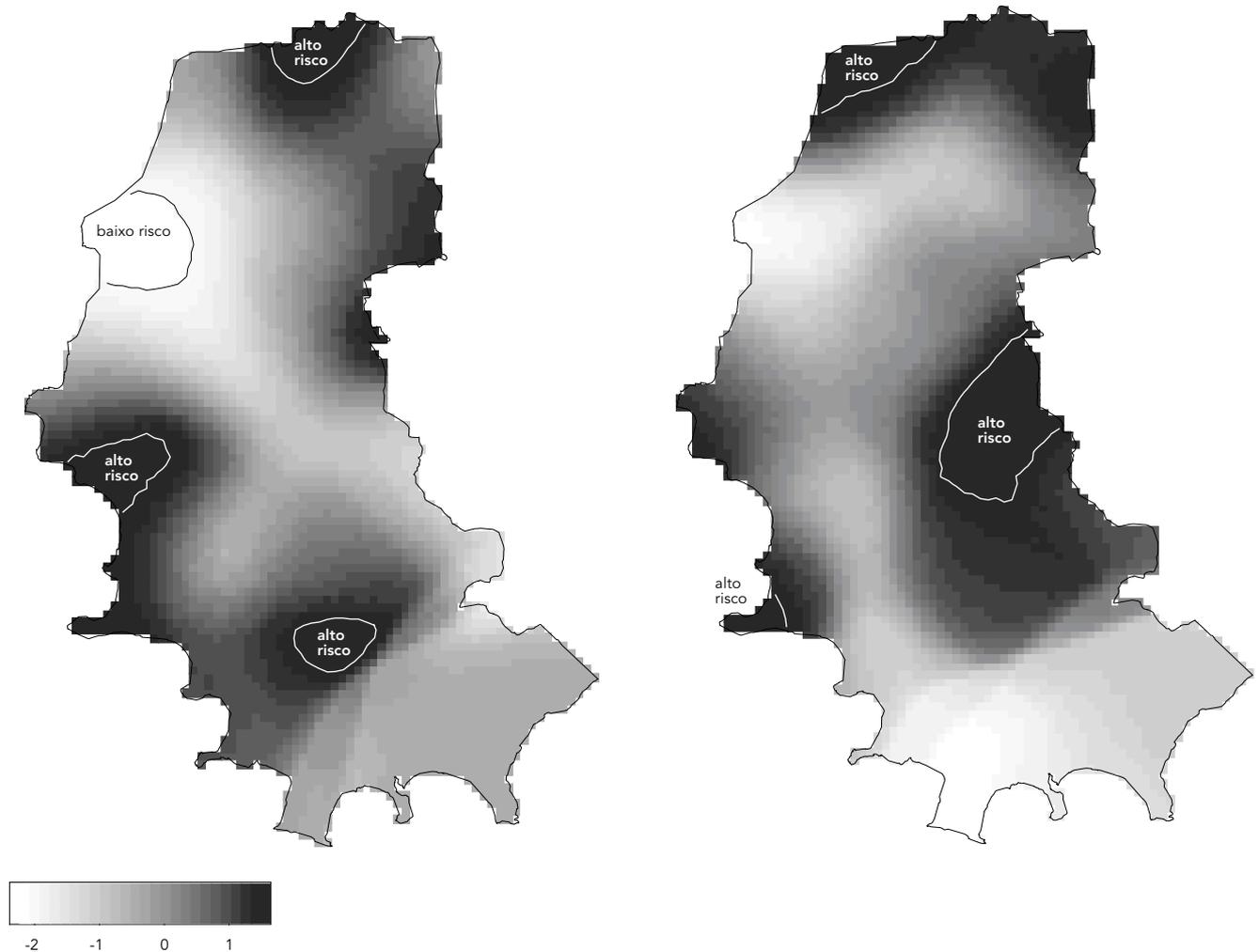
Em relação aos dois valores de banda utilizados, ambos permitiram identificar variação espacial significativa. A vantagem do método de validação cruzada é que independe de critérios subjetivos. E, possivelmente, por este motivo, a largura de banda de 3.000m, maior, portanto, que a escolhida empiricamente, obtida para a mortalidade neonatal, permitiu identifi-

car somente uma área de alto risco, que coincide com o esperado, baseado no conhecimento prévio das condições de vida da população. Entretanto, sendo o alisamento Kernel muito flexível e com grande potencial para análise exploratória, a interferência do pesquisador na seleção da largura de banda, incluindo o conhecimento prévio da área e do problema em estudo, pode ser interessante.

Um desdobramento importante seria a comparação dessas superfícies de risco com mapas nos quais se pudessem detectar particularidades no perfil de carência social das regiões de sobre-risco. Esses achados, portanto, não são conclusivos; muito pelo contrário, apenas fornecem mais elementos para futuras investigações. A não identificação de áreas de sobre-risco na mortalidade pós-neonatal pode estar relacionada ao acesso espacialmente homogêneo a serviços da rede básica de saúde que atendem às principais causas de morte nesta faixa etária. É possível que o contrário esteja acontecendo no caso da mortalidade neonatal: diferenças na distribuição espacial dos serviços de saúde e do acesso a recursos de assistência ao pré-natal e ao parto estariam condicionando regiões com maior mortalidade neonatal. Além disso, o menor número de obser-

Figura 2

Mapas de risco para a mortalidade infantil, controlando para fatores individuais – largura de banda = 2km. Porto Alegre, 1998.

Figura 2a: Neonatal ($p = 0,024$)Figura 2b: Pós-neonatal ($p = 0,137$)

vações de óbitos pós-neonatais também pode contribuir para a não significância.

A característica de ferramenta exploratória é, de fato, marcante dos métodos estatísticos de mapeamento espacial do risco. O resultado da análise indica claramente os territórios em maior risco na cidade, possibilitando o planejamento de ações mais específicas e dirigidas a essas áreas em um contexto de vigilância da saúde, cuja unidade de entendimento, planeja-

mento e ação é o território e seus múltiplos recortes.

Os determinantes de risco da mortalidade infantil aqui identificados são mediadores entre o processo de desenvolvimento brasileiro e as condições de vida e saúde das crianças. É por meio da ação desses determinantes que os processos estruturais da sociedade deixam marcas na saúde dessas crianças, muitas vezes tão decisivas que provocam seu óbito.

Figura 3

Mapas de risco para a mortalidade infantil, controlando para fatores individuais – largura de banda estimada por validação cruzada. Porto Alegre, 1998.

Figura 3a: Neonatal - cv (p = 0,012)

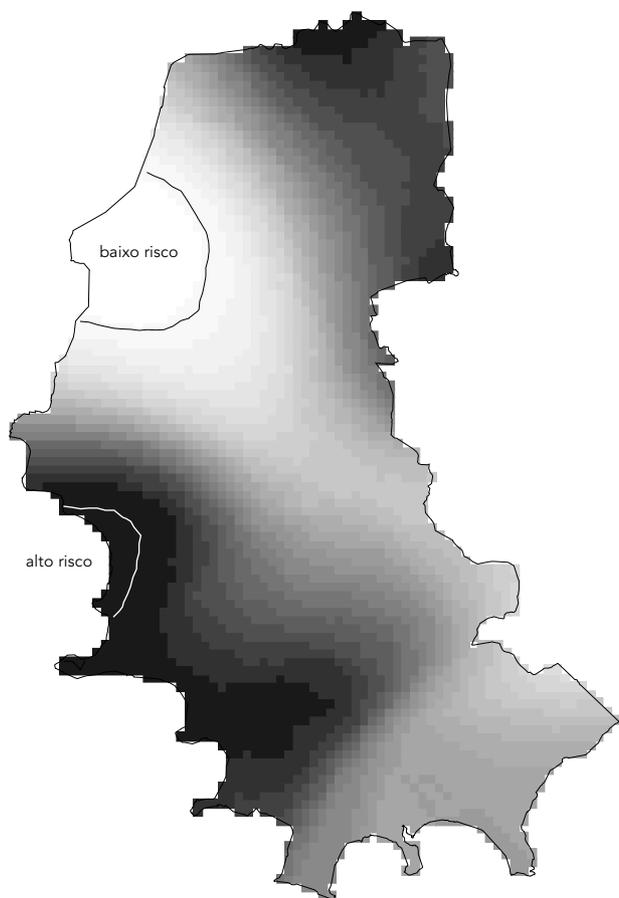
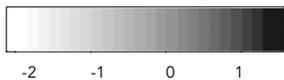
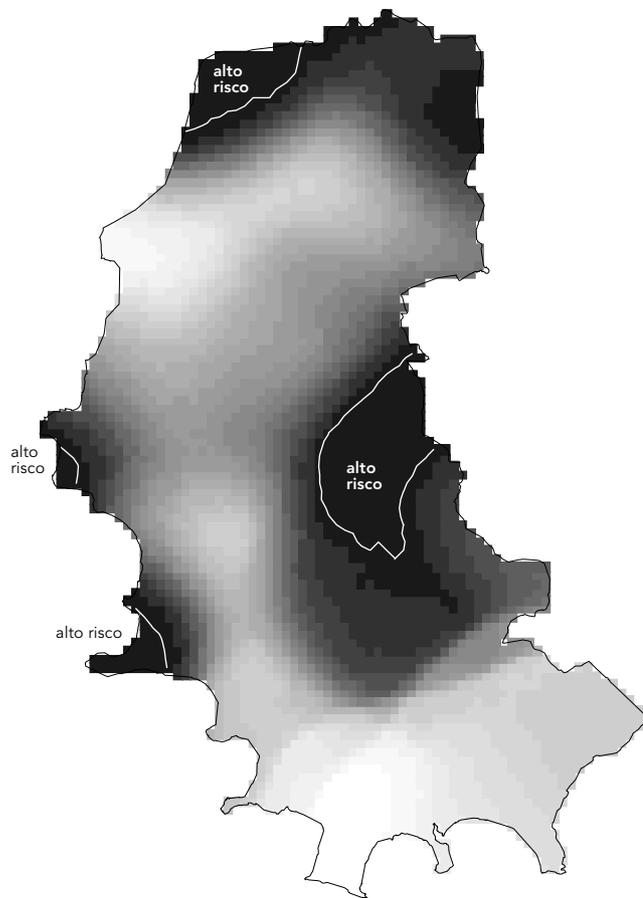


Figura 3b: Pós-neonatal - cv (p = 0,084)



Referências

- AERTS, D., 1992. *Estudo do Estado Nutricional das Crianças de Porto Alegre: Uma Contribuição ao Entendimento do Processo da Desnutrição*. Dissertação de Mestrado, Porto Alegre: Universidade Federal do Rio Grande do Sul.
- AERTS, D., 1996. *O Retardo no Crescimento e seus Determinantes: O Caso de Porto Alegre*. Tese de Doutorado, Porto Alegre: Universidade Federal do Rio Grande do Sul.
- BARCELLOS, T.; BARROS, C.; OLIVEIRA, N.; GARCIA, V.; PASTORE, L. & RABELO, M., 1986. *Segregação Urbana e Mortalidade em Porto Alegre*. Porto Alegre: Fundação de Economia e Estatística Siegfried Emanuel Heuser.
- BITHELL, J. F., 1992. Statistical methods for analysing point-source exposures. In: *Geographical and Environmental Epidemiology: Methods for Small Area Studies* (P. Elliott, P. Cuzick, J. English & D. Stern, ed.), pp. 221-230, Oxford: Oxford University Press.
- CAMPOS, T. P.; CARVALHO, M. S. & BARCELLOS, C., 2000. Mortalidade infantil no Rio de Janeiro: Áreas de risco e trajetória dos pacientes aos serviços. *Revista Panamericana de Salud Publica*, 8:164-171.
- DIEZ-ROUX, A. V., 1998. Bringing context back into epidemiology: Variables and fallacies in multi-level analysis. *American Journal of Epidemiology*, 88:216-222
- DIGGLE, P. J., 1983. *Statistical Analysis of Spatial Point Patterns*. London: Academic Press.
- DIGGLE, P. J., 1992. *Point Process Modelling in Environmental Epidemiology*. Technical Report MA92/70. Lancaster: Department of Mathematics and Statistics, Lancaster University.
- FAN, J. Q.; HECKMAN, N. E. & WAND, M. P., 1995. Local polynomial kernel regression for generalized linear models and quasi-likelihood functions. *Journal of the American Statistical Association*, 90:141-150.
- GUIMARÃES, J. J. & FISCHMANN, A., 1986. Desigualdades na mortalidade infantil entre favelados e não favelados no município de Porto Alegre, Rio Grande do Sul, em 1980. *Boletín de la Oficina Sanitaria Panamericana*, 101:19-39.
- HÄRDLE, W., 1990. *Applied Nonparametric Regression*. Cambridge: Cambridge University Press.
- HASTIE, T. J. & TIBSHIRANI, R. J., 1990. *Generalized Additive Models*. London: Chapman and Hall.
- HOSMER, D. W. & LEMESHOW, S., 1989. *Applied Logistic Regression*. New York: Wiley.
- KELSALL, J. E., 1996. *Kernel Smoothing for Application in Environmental Epidemiology*. Ph.D. Thesis, Lancaster: Department of Mathematics and Statistics, Lancaster University.
- KELSALL, J. E. & DIGGLE, P. J., 1995a. Kernel estimation of relative risk. *Bernoulli*, 1:3-16.
- KELSALL, J. E. & DIGGLE, P. J., 1995b. Non-parametric estimation of spatial variation in relative risk. *Statistics in Medicine*, 14:2335-2342.
- KELSALL, J. E. & DIGGLE, P. J., 1998. Spatial variation in risk of disease: A nonparametric binary regression approach. *Applied Statistics*, 47:559-573.
- LAWSON, A. B. & WILLIAMS, F. L. R., 1993. Applications of extraction mapping in environmental epidemiology. *Statistics in Medicine*, 12:1249-1258.
- MARTINS, C. & LOUREIRO, A., 1996. Porto Alegre e os anos 90: Características sócio-econômicas da metrópole gaúcha. In: *Dinâmica da Urbanização no Rio Grande do Sul: Temas e Tendências* (N. Oliveira, T. Barcellos, C. Martins, A. Breitbart, A. Loureiro & S. Borba, org.), pp. 152-203, Porto Alegre: Fundação de Economia e Estatística Siegfried Emanuel Heuser.
- MATHSOFT INCORPORATION, 1998. *S-Plus for Windows, Version 5.5*. Seattle: Mathsoft Inc.
- OLIVEIRA, L. A. P. & SIMÕES, C. C. S., 1986. O papel das políticas de saneamento na recente queda da mortalidade infantil: Significado, alcance e limitações estruturais. In: *Perfil Estatístico de Crianças e Mães no Brasil: Aspectos Sócio-econômicos da Mortalidade Infantil em Áreas Urbanas*. (Fundação Instituto Brasileiro de Geografia e Estatística & Fundo das Nações Unidas para a Infância, org.), pp. 51-57, Rio de Janeiro: IBGE.
- SMSSS (SECRETARIA MUNICIPAL DE SAÚDE E SERVIÇO SOCIAL), 1989. *Plano Municipal de Saúde*. Porto Alegre: SMSSS. (mimeo.)
- WAND, M. P. & JONES, M. C., 1995. *Kernel Smoothing*. London: Chapman and Hall.