



JOSÉ VILTON COSTA

**ANÁLISE ESPACIAL DO RISCO DE DENGUE NO
MUNICÍPIO DE CAMPINAS: MODELAGEM BAYESIANA**

**CAMPINAS
2013**



UNIVERSIDADE ESTADUAL DE CAMPINAS
FACULDADE DE CIÊNCIAS MÉDICAS

JOSÉ VILTON COSTA

**ANÁLISE ESPACIAL DO RISCO DE DENGUE NO
MUNICÍPIO DE CAMPINAS: MODELAGEM BAYESIANA**

Orientador(a): Prof(a). Dr(a). Liciana Vaz de Arruda Silveira

Co-Orientador(a): Prof(a). Dr(a). Maria Rita Donalísio Cordeiro

Tese de Doutorado apresentada ao Programa de Pós-Graduação em Saúde Coletiva da Faculdade de Ciências Médicas da Universidade de Campinas para obtenção de título de Doutor em Saúde Coletiva, área de concentração Epidemiologia.

ESTE EXEMPLAR CORRESPONDE À VERSÃO FINAL DA TESE
DEFENDIDA PELO ALUNO JOSÉ VILTON COSTA E
ORIENTADO PELO PROF^ª. DR^ª. LICIANA VAZ DE ARRUDA
SILVEIRA.

CAMPINAS
2013

FICHA CATALOGRÁFICA ELABORADA POR
MARISTELLA SOARES DOS SANTOS – CRB8/8402
BIBLIOTECA DA FACULDADE DE CIÊNCIAS MÉDICAS
UNICAMP

C823a Costa, José Vilton, 1973-
Análise espacial do risco de dengue no município de
Campinas : modelagem bayesiana / José Vilton Costa . --
Campinas, SP : [s.n.], 2013.

Orientador : Liciano Vaz de Arruda Silveira.
Coorientador : Maria Rita Donalísio Cordeiro.
Tese (Doutorado) - Universidade Estadual de
Campinas, Faculdade de Ciências Médicas.

1. Dengue. 2. Análise espacial. 3. Epidemiologia. 4.
Estudos ecológicos. 5. Teoria bayesiana de decisão
estatística. I. Silveira, Liciano Vaz de Arruda. II.
Cordeiro, Maria Rita Donalísio. III. Universidade
Estadual de Campinas. Faculdade de Ciências Médicas.
IV. Título.

Informações para Biblioteca Digital

Título em inglês: Spatial analysis of risk of dengue in the municipality of Campinas :
bayesian modeling.

Palavras-chave em inglês:

Dengue

Spatial analysis

Epidemiology

Ecological studies

Bayesian statistical decision theory

Área de concentração: Epidemiologia

Titulação: Doutor em Saúde Coletiva

Banca examinadora:

Liciano Vaz de Arruda Silveira [Orientador]

Ricardo Carlos Cordeiro

Celso Stephan

José Silvío Govone

Lidia Raquel de Carvalho

Data da defesa: 01-02-2013

Programa de Pós-Graduação: Saúde Coletiva

BANCA EXAMINADORA DA DEFESA DE DOUTORADO

JOSÉ VILTON COSTA

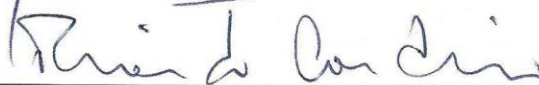
Orientador (a) PROF(A). DR(A). LICIANA VAZ DE ARRUDA SILVEIRA

MEMBROS:

1. PROF(A). DR(A). LICIANA VAZ DE ARRUDA SILVEIRA



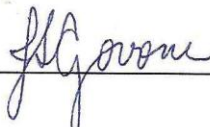
2. PROF(A). DR(A). RICARDO CARLOS CORDEIRO



3. PROF(A). DR(A). CELSO STEPHAN



4. PROF(A).DR(A). JOSÉ SILVIO GOVONE



5. PROF(A).DR(A). LIDIA RAQUEL DE CARVALHO



Programa de Pós-Graduação em Saúde Coletiva da Faculdade de Ciências Médicas da
Universidade Estadual de Campinas

Data: 01 de fevereiro de 2013

AGRADECIMENTOS

Dedico um agradecimento especial as minhas orientadoras, a Profa. Liciane e Profa. Rita, que durante todo o período em que este trabalho foi desenvolvido me orientaram, apoiaram, e sempre estiveram presentes e dedicaram muita atenção para que tudo fosse resolvido da melhor forma possível.

À CAPES pela concessão da bolsa de estudos.

Aos meus professores do Programa de Pós-Graduação em Saúde Coletiva por compartilharem conosco o seu conhecimento e experiência, contribuindo com o meu desenvolvimento científico.

À bióloga Heloísa Malavasi pelo excelente trabalho realizado no georreferenciamento dos dados de dengue utilizados nesta tese.

A todos os meus colegas da Pós.

Aos membros da banca de qualificação, prof. Ricardo e prof. Celso, que contribuíram com suas sugestões e incentivo. E aos novos membros da banca de defesa, a profa. Lídia e o prof. Sílvio.

À minha esposa Verena, aos meus filhos Tiago e Isadora, pelo apoio, compreensão, amor e estímulo ao longo desse trabalho.

Aos meus pais pelo imenso esforço realizado para que eu pudesse ter uma educação de qualidade.

Ao meus irmão e irmãs pela amizade e incentivo.

Aos meu sogros por toda a ajuda e apoio ao longo desta jornada.

RESUMO

A dengue é uma doença viral de transmissão vetorial causada por um dos quatro sorotipos do vírus dengue (DENV-1, DENV-2, DENV-3 e DENV-4). Atualmente é considerada um dos maiores problemas de Saúde Pública do Mundo. O controle do vetor transmissor, *Aedes aegypti*, é bem complexo, sendo de grande importância para os serviços de vigilância epidemiológica compreender o processo espacial da dinâmica de transmissão da doença sobre o espaço intramunicipal. A presente tese tem como objetivo analisar a distribuição espacial do risco de dengue e sua associação com fatores socioambientais, no município de Campinas-SP, em 2007. Considerando-se a escala local, analisou-se duas distintas dimensões espaciais: os setores censitários e as áreas de cobertura dos centros de saúde. Foi construído um indicador de condições socioambientais a partir de variáveis do Censo 2000. Foram investigados 11519 casos de dengue autóctones. Na modelagem de regressão ecológica, foram incluídos efeitos aleatórios estruturados espacialmente e não estruturados para ajustar a variação extra-Poisson presente nas contagens agregadas da doença. Os modelos binomial negativo (BN), Poisson inflacionado de zeros (ZIP) e binomial negativo inflacionado de zeros (ZINB) foram aplicados para modelar a superdispersão e o excesso de zeros para os dados dos setores censitários. A abordagem hierárquico bayesiana foi utilizada para inferência e o método INLA (*Integrated Nested Laplace Approximations*) foi empregado para estimação dos parâmetros. A distribuição espacial do risco de dengue sobre o espaço intramunicipal de Campinas não mostrou clara relação com as condições socioambientais. O modelo ZIP mostrou-se mais adequado para modelar o excesso de zeros presente nos dados de contagem para pequenas áreas.

ABSTRACT

Dengue is a mosquito-borne disease caused by four serotypes of dengue virus (DENV1 to DENV4) and is currently one of the major public health problems worldwide. The control of its vector, the *Aedes aegypti* mosquito, is very complex and it is of great importance for epidemiological surveillance systems to understand the process of spatial dynamics of disease transmission on urban space. This thesis aimed to analyze the spatial distribution of the risk of dengue and its association with socio-environmental factors, in Campinas-SP, in 2007. Considering the local scale, we analyzed two distinct spatial dimensions: census tracts and health districts. We constructed an indicator of socio-environmental conditions from 2000 Census data. We investigated 11,519 cases of autochthonous dengue. In ecological modeling regression, random effects were included spatially structured and unstructured to fit the extra-Poisson variation present in the count data of disease. The negative binomial models (BN), zero inflated Poisson (ZIP) and zero inflated negative binomial (ZINB) were applied to model the overdispersion and excess zeros to the data of the census tracts. A hierarchical Bayesian approach was used for inference method and the INLA (Integrated Nested Laplace approximations) was used for parameter estimation. The spatial distribution of the risk of dengue on local scale Campinas showed no clear relationship with the socio-environmental conditions. The ZIP model was more appropriate for modeling excess zeros count data present in small areas.

LISTA DE ABREVIATURAS E SIGLAS

ACCS	Áreas de Cobertura dos Centros de Saúde
ACP	Análise de Componentes Principais
BN	Binomial negativa
CAMG	Campo aleatório Markoviano Gaussiano
CAR	Conditional Autorregressive
Covisa	Coordenadoria de Vigilância e Saúde Ambiental
CPO	Conditional predictive ordinate (Ordenada preditivacondicional)
DC	Dengue clássica
DENV-1	Dengue virus sorotipo 1
DENV-2	Dengue virus sorotipo 2
DENV-3	Dengue virus sorotipo 3
DENV-4	Dengue virus sorotipo 4
DIC	Deviance Information Criteria (critério de informação da deviance)
FHD	Febre hemorrágica de dengue
GLM	Generalized Linear Model (modelos lineares generalizados)
IBGE	Instituto Brasileiro de Geografia e Estatística
IC	Intervalo de credibilidade
ICSA	Índice de condições socioambientais
INLA	Integrated Nested Laplace <i>Approximations</i>
MCMC	Markov chain Monte Carlo (Monte Carlo em cadeias de Markov)
OMS	Organização Mundial de Saúde
OPAS	Organização Pan Americana de Saúde
RR	Risco relativo
SCD	Síndrome de choque da dengue
SINAN	Sistema de Informação de Agravos de Notificação
SIR	Standardized Incidence Ratio
SMS	Secretaria Municipal de Saúde
STAR	Structured additive regression
ZI	Zero-inflated (inflacionado de zeros)
ZINB	Zero-inflated negative binomial (binomial negativa inflacionada de zeros)
ZIP	Zero-inflated Poisson (Poisson inflacionado de zeros)

LISTA DE ILUSTRAÇÕES

Figura 1. Distribuição Global do risco de dengue.....	14
Figura 2. Número médio de casos de dengue registrados pela OMS nos 30 países mais endêmicos, 2004-2010.....	17
Figura 3. Distribuição dos casos autóctones de dengue por início dos sintomas, Estado de São Paulo.....	21

SUMÁRIO

1. INTRODUÇÃO.....	12
1.1. Dengue	13
1.2. Dengue nas Américas.....	15
1.3. Dengue no Brasil.....	18
1.4. Dengue no Estado de São Paulo.....	19
1.5. Espaço e Dengue	21
1.6. Análise de dados espaciais em saúde	24
1.7. Modelos Inferenciais Bayesianos.....	26
2. OBJETIVOS.....	28
2.1. Objetivo Geral	29
2.2. Objetivos Específicos	29
3. MATERIAIS E MÉTODOS	30
3.1. Desenho do estudo.....	31
3.2. Área de estudo	31
3.3. Dados.....	31
3.4. Estimaco do risco relativo para dados agregados	33
3.5. Modelos Hierárquicos Bayesianos	38
3.6. Efeitos aleatrios	39
3.7. Método INLA	42
3.8. Comparaco dos modelos bayesianos	43
4. RESULTADOS	45
4.1. Artigo 1	46
4.2. Artigo 2	66
5. DISCUSSO E CONCLUSO GERAL	88
6. REFERNCIAS BIBLIOGRÁFICAS.....	93
7. ANEXOS.....	100
7.1. Anexo 1. Comprovante de aceite do artigo submetido ao Cad. Sade Publica.....	100

1. INTRODUÇÃO

1.1. Dengue

Dengue é uma doença viral de transmissão vetorial causada por um dos quatro sorotipos do vírus dengue (DENV-1, DENV-2, DENV3 e DENV-4), pertencentes ao gênero *Flavivirus*, da família *Flaviviridae*(1). É endêmica de regiões tropicais como o sudeste asiático, sul do Pacífico, África Oriental, Caribe e América Latina. Atualmente vem preocupando as autoridades em saúde em quase todo o mundo, devido à sua ampla distribuição e ao grande potencial para causar casos graves e letais (2).

A Organização Mundial da Saúde (OMS) (3), considera a dengue um dos maiores problemas de saúde pública, com cerca de 100 milhões de casos anualmente no mundo, sendo que mais de 2,5 bilhões de pessoas vivem em área sob risco de infecção (Figura 1).

É uma doença febril, que pode ser de curso benigno ou grave. Dependendo de suas manifestações recebe a seguinte tipologia: infecção inaparente, dengue clássico (DC), febre hemorrágica da dengue (FHD) ou síndrome do choque do dengue (SCD) (1).

O vírus é transmitido ao homem pela picada do mosquito fêmea infectado do gênero *Aedes*, em particular pela espécie *Aedes aegypti* (1). O *Aedes aegypti*, é uma das mais de 500 espécies de *Aedes* (*Diptera: Culicidae*), gênero com ampla distribuição e adaptabilidade. É considerada uma espécie autóctone do continente africano e acredita-se que a Etiópia tenha sido o centro da dispersão.

A transmissão dos vírus de dengue pode ocorrer imediatamente após a troca de hospedeiros ou depois de um período de incubação, no qual há multiplicação viral (4). O vetor põe seus ovos nas folhas de plantas que armazenam água (p.ex. bromélias) ou nas paredes de recipientes, como pneus, vasos e pias, entre outros. São locais que se constituem em criadouros naturais e artificiais, respectivamente, sendo esses últimos os mais importantes do ponto de vista epidemiológico (5).

A imunidade conferida pela infecção é permanente para o sorotipo que a causou (imunidade homóloga), mas temporária e parcial para os outros sorotipos (heteróloga). A patogênese da FHD não está totalmente esclarecida, várias teorias tentam explicar a sua ocorrência. Segundo a OMS, duas infecções sequenciais por diferentes sorotipos podem

provocar uma resposta imunológica exacerbada na segunda infecção, desencadeando essa forma clínica da doença(1).

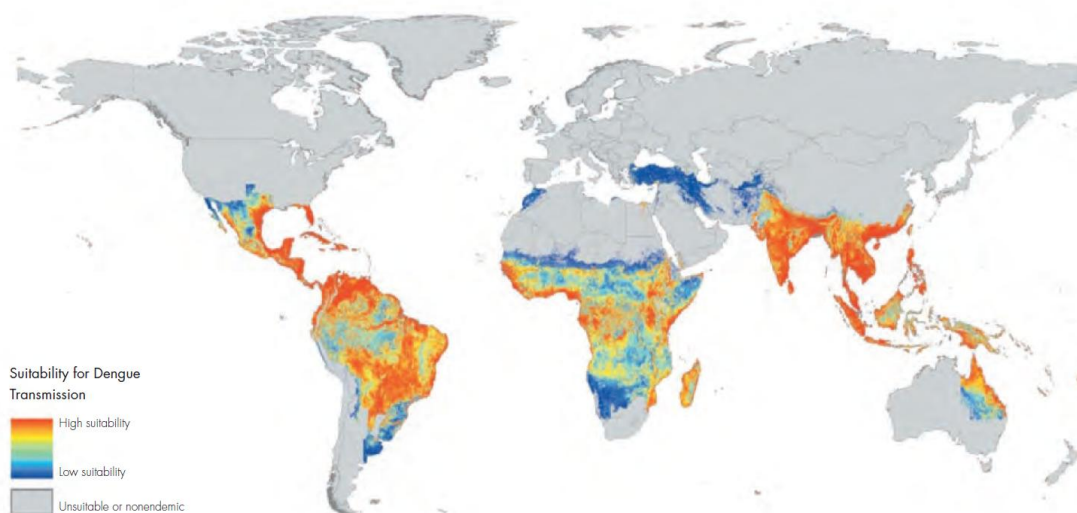


Figura 1. Distribuição Global do risco de dengue.

Extraído de: OMS (3).

Os mosquitos se desenvolvem através de metamorfose completa e o ciclo de vida do *Aedes aegypti* compreende quatro fases: ovo, larva (quatro estágios larvários), pupa e adulto (4). Os ovos, que podem resistir por vários meses à dessecação, eclodem após a submersão. Ao longo da vida, de aproximadamente dois meses, uma única fêmea de *Aedes aegypti* é capaz de transmitir os vírus da dengue até 12 vezes ou mais (6,7).

O mosquito possui hábitos domésticos e a grande variedade de criadouros potenciais encontrados no peridomicílio e no intradomicílio garante a manutenção de altas densidades de *Aedes aegypti* no meio urbano, potencializando os riscos de transmissão da doença (1). Uma vez que a dispersão de dengue se processa principalmente de domicílio a domicílio e de modo contínuo, em geral, epidemias de dengue nas áreas urbanas são explosivas e envolvem porções apreciáveis da população, especialmente durante as estações chuvosas quando existe uma maior abundância do vetor (1,6).

Atualmente não se dispõe de uma vacina eficaz para uso preventivo contra dengue (8). Embora, um estudo com uma vacina candidata a tetravalente tenha demonstrado que uma vacina segura e eficaz é possível (9). Na ausência de uma vacina

eficaz, as estratégias de prevenção e controle são direcionadas ao combate direto ao mosquito vetor, considerado o único elemento controlável da cadeia epidemiológica da dengue (10). As principais estratégias de controle estão direcionadas a vigilância vetorial e no controle direto ao vetor por combate químico, físico e biológico (11).

No entanto, os métodos de controle direto têm apresentado baixa efetividade e altos custos. Diante disto e da gravidade da doença, tem sido discutido e adotadas outras estratégias de prevenção, baseadas na participação comunitária, educação em saúde e na melhoria das condições socioambientais da população (12).

A dinâmica e diferenciação na transmissão do vírus da dengue de um lugar a outro está determinada pelas interações entre o ambiente, o vetor, o agente etiológico e a população de hospedeiros suscetíveis, que co-existem juntos num habitat específico. A magnitude e intensidade destas interações definem a transmissão em uma localidade (13,14).

Os fatores de risco para transmissão da dengue estão relacionados aos chamados macrofatores e microfatores. Entre os macrofatores se encontram os fatores de risco sociais e ambientais e entre os microfatores estão as características biológicas do vírus, do vetor e do hospedeiro (15).

1.2. Dengue nas Américas

Nas Américas, os primeiros relatos clínicos e epidemiológicos condizentes com a de dengue ocorreram na Filadélfia, em 1790. Alguns anos mais tarde, em 1827, houve uma epidemia no Caribe. Em 1881-1885 reapareceu no Texas, foi relatada em Cuba e, novamente, no Texas, em 1897. Apareceu em Porto Rico em 1915 e Jamaica em 1917, espalhando-se por todo o Caribe em 1922. A amplitude no intervalo entre as epidemias e flutuação geográfica impediu que a dengue fosse identificado como um problema de saúde em cada país (16).

A Segunda Grande Guerra amplificou a população do *Aedes aegypti* e a distribuição do vírus dengue no Sudeste Asiático. Nas Américas, diferentemente, o programa de controle da Febre Amarela organizado pela Organização Panamericana de Saúde (OPS) nas

décadas de 50 e 60 levou a erradicação do *Aedes Aegypti* da maioria dos países da América Central e do Sul. Porém, o programa de erradicação do mosquito transmissor foi oficialmente descontinuado pelos Estados Unidos em 1970, levando a reinfestação dos países de onde o mosquito já havia sido erradicado (17,18).

De 1963 a 1977, alguns surtos associados aos DENV-2 e 3 foram confirmados no Caribe, Haiti, República Dominicana, Porto Rico e norte da América do Sul. A re-emergência da dengue nas Américas foi marcada pela introdução do sorotipo DENV-1, em 1977, seguida de uma devastadora epidemia que durou até 1980. Originalmente, o vírus foi detectado na Jamaica, expandindo-se por países da América Central e do Sul, e posteriormente para México e Estados Unidos. Cerca de 702.000 casos de dengue foram notificados de 1977 a 1980, período no qual o DENV-1 foi praticamente o único sorotipo circulante (19).

No ano de 1981, o DENV-4 surgiu nas Américas, provavelmente importado das ilhas do Pacífico, o que causou uma série de surtos no Caribe, norte da América do Sul, América Central e México. Durante os anos oitenta cinco países sul-americanos (Brasil, Bolívia, Paraguai, Equador e Peru) que não tinham tido dengue antes ou que tinha sido livre da doença por várias décadas foram afetados por epidemias explosivas causadas pelo sorotipo DENV-1 (20).

A década de 80 foi marcada pela co-circulação dos DENV-1, DENV-2 e DENV-4, e países como Brasil, Bolívia, Paraguai, Equador e Peru, que até então não haviam experimentado uma epidemia de dengue ou estiveram livres da doença nas décadas anteriores, foram acometidos por grandes epidemias (19,20).

Em 1993, os últimos países tropicais da América Latina, Costa Rica e Panamá, livres de dengue durante várias décadas, relataram casos de transmissão de dengue autóctone, do sorotipo DENV-1. Em 1994 o DENV-3 foi reintroduzido nas Américas após 16 anos de ausência. Esse sorotipo foi inicialmente detectado no Panamá, Nicarágua e posteriormente em outros países da América Central e México, causando numerosas epidemias. No México, a introdução do DENV-3 em 1995 foi associada com a gravidade da doença (19).

A situação epidemiológica da dengue nas Américas tem se agravado durante os últimos 20 anos, demonstrando-se que a dengue é endêmica nas Américas. Nota-se na

região um aumento no número de casos e de países atingidos, com a ocorrência de picos epidêmicos cada vez maiores que se repetem ciclicamente no período de 3 a 5 anos (21). A maior incidência na região foi observada em 2002, com mais de um milhão de casos, entre eles 14.000 de DH. Dados da OMS sobre a incidência da doença (3), no período 2004-2010, demonstram que entre os 30 países do mundo com maior incidência de dengue, 18 destes países são das Américas, em que se destaca o Brasil por apresentar o maior número de casos da doença (Figura 2).

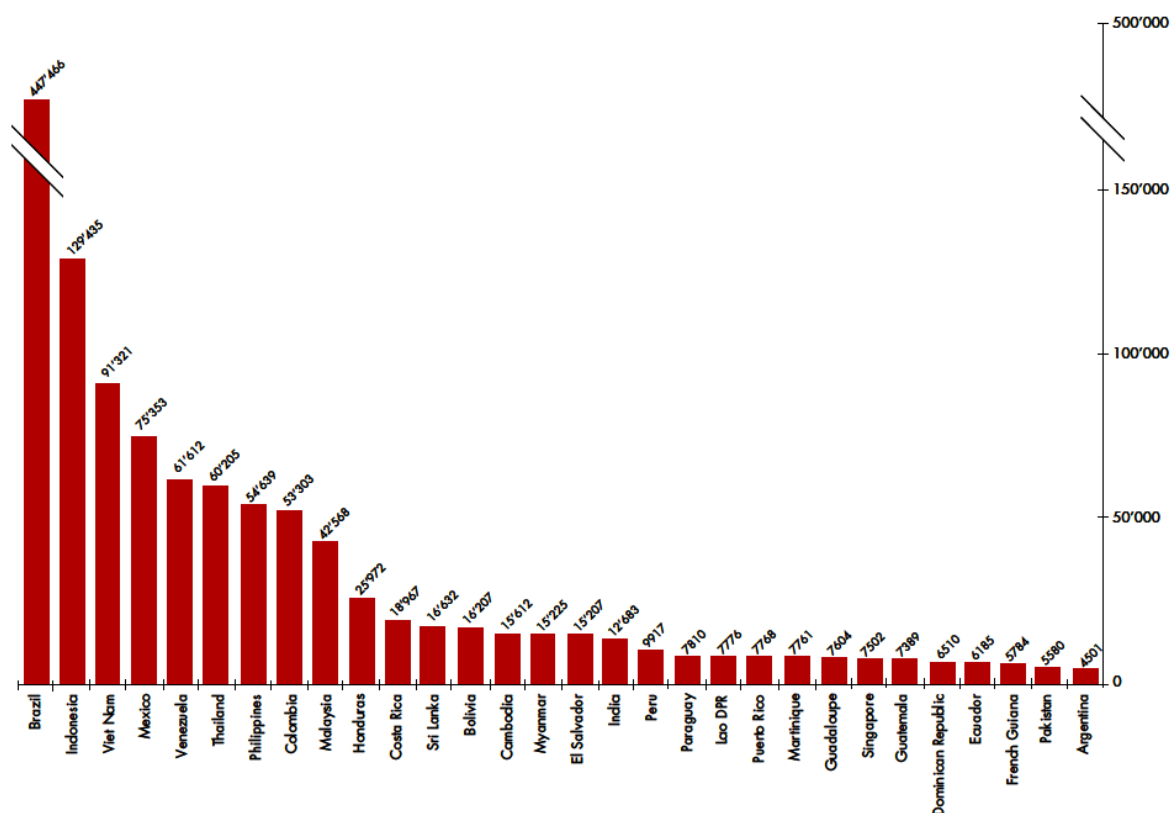


Figura 2. Número médio de casos de dengue registrados pela OMS nos 30 países mais endêmicos, 2004-2010.

Extraído de: OMS (3).

1.3. Dengue no Brasil

A reinfestação do *Aedes aegypti* no Brasil, a partir do biênio 1976-1977, criou o elo básico da cadeia epidemiológica para a reintrodução da dengue no país (1). A primeira epidemia, documentada clínica e laboratorialmente, ocorreu no início da década de 1980, em Boa Vista, Estado de Roraima, com a notificação de 7.000 casos causados pelos sorotipos DENV-1 e DENV-4 (22). Após um silêncio epidemiológico, em 1986, o DENV-1 foi isolado na cidade de Nova Iguaçu, Rio de Janeiro (23). A partir daí o vírus rapidamente se espalhou por outras cidades, e a dificuldade em implantar programas efetivos de combate ao vetor resultou na dispersão do DENV-1 para os estados do Nordeste e para sua interiorização, alcançando São Paulo, Minas, Mato Grosso do Sul e, posteriormente, outros estados, incluindo Goiás na década de 1990 (24).

No biênio 1990/1991, a situação epidemiológica se agravou com a entrada do sorotipo DENV-2 no Brasil, com 104.398 casos no final deste biênio. Em 1990, somente no Estado do Rio de Janeiro, a incidência da doença atingiu 165,7 casos por 100 mil habitantes, e, em 1991, 613,8 casos por 100 mil habitantes. Nesse mesmo período, os primeiros casos de FHD são registrados no país (7,25). Em sequência, os vírus DENV-1 e DENV-2 foram se disseminando para outros Estados do país já infestados pelo *Aedes aegypti*, aumentando o número casos de dengue e a ocorrência de epidemias (7,24). Epidemias de dengue causadas pelo DENV-2 ocorreram no Tocantins em 1998 e no Estado do Pará em 1995. Em 1998, no Amazonas, foi registrado uma epidemia com 23.910 casos de dengue (22).

Em 2001, foi confirmada no Rio de Janeiro a introdução no país do sorotipo DENV-3, produzindo novamente uma epidemia em larga escala epidemiológica, expandido rapidamente e afetando pequenas cidades que até então estavam livres da dengue. Esse sorotipo foi responsável pela epidemia de 2002 no Brasil, quando foram notificados aproximadamente 800 mil casos, ou seja, quase 80% das ocorrências do continente americano. A introdução e predomínio do DENV-3 marca o aumento da ocorrência de formas graves na doença no país (24).

Além da ocorrência de epidemias de grande magnitude, a alternância de sorotipos predominantes tem levado a importantes alterações na epidemiologia da doença.

No início da década de 2000, com o predomínio do sorotipo DENV-3, houve aumento da proporção de casos graves. Ao final da década, as epidemias causadas pelo DENV-2 foram marcadas pelo aumento de casos graves entre crianças. Em 2010, ocorreu a maior epidemia de dengue já registrada no Brasil, com predomínio do sorotipo DENV-1 (26).

No Brasil, o padrão de transmissão de dengue é caracterizado por ciclos de predomínio de um determinado sorotipo do vírus. Na década de 2000, cada ciclo foi caracterizado por novos períodos de alta transmissão da doença após cada mudança. Tem-se observado importantes alterações na epidemiologia da doença em decorrência de epidemias de grande magnitude e da alternância de sorotipos predominantes. As epidemias de DENV-3 no início da década levaram a um aumento na gravidade dos casos, enquanto as epidemias causadas pelo DENV-2 no final da década foram marcadas pelo aumento de casos graves em crianças. Ao final do ano de 2009 teve início um novo ciclo de alta transmissão do DENV-1, com o registro de mais de um milhão de casos prováveis no ano de 2010. As epidemias desse ano afetaram principalmente áreas mais ao centro do país, culminando no maior número de óbitos pela doença em um único ano até o momento no país (26).

No segundo semestre de 2010 foram confirmados os primeiros casos de dengue pelo sorotipo DENV-4 na região Norte (em Roraima, Amazonas e Pará). Em poucos meses a circulação desse sorotipo foi detectada também no Nordeste (Piauí e Bahia) e Sudeste (Rio de Janeiro). A introdução deste sorotipo DENV-4 criou a possibilidade de ocorrência de novas epidemias no país (26). Embora este sorotipo ainda não tenha se tornado predominante até o momento (27), diferentes regiões do mundo apresentaram a rápida substituição de outros sorotipos pelo DENV-4 (28,29).

1.4. Dengue no Estado de São Paulo

No Estado de São Paulo, a transmissão de dengue autóctone, confirmada laboratorialmente, foi detectada pela primeira vez em 1987, nos municípios de Guararapes (30 casos) e Araçatuba (16 casos). Casos confirmados de dengue já haviam sido verificados no Estado no ano anterior, porém todos importados, sendo a maioria do Rio de Janeiro. Entre os anos de 1988 e 1990, houve registros de casos, no entanto, nenhum autóctone do Estado (30).

No verão de 1990/91 foi registrada uma epidemia de grandes proporções, com início em Ribeirão Preto, que rapidamente se expandiu para municípios vizinhos e outras regiões, com a confirmação de 6.701 casos em 59 municípios. A partir de então, as epidemias de dengue vêm ocorrendo todos os anos no Estado (30).

Até 1996, as epidemias eram devidas ao sorotipo DENV-1. A partir desse ano detectou-se a introdução do DENV-2 (30). Em 2002, foi registrada no Estado a introdução do sorotipo DENV-3, sendo identificado em 20 municípios. Com a sucessiva exposição da população paulista aos diferentes sorotipos, o risco das formas graves da doença aumentou (31).

O ano de 1997 marca a expansão da transmissão de dengue para a região da Baixada Santista, constituindo-se nos anos seguintes a região mais importante de transmissão, atingindo seu ápice em 2002, quando respondeu por 69% dos casos do Estado; e na Grande São Paulo em 2002 com elevada incidência já em 2003, sendo responsável por 44% da transmissão do Estado. O primeiro caso de dengue hemorrágico foi confirmado em 1999. Nos anos seguintes: 2, 5, 31 e 20 casos, respectivamente em 2000, 2001, 2002 e 2003.

O Estado de São Paulo, nos últimos 10 anos, tem vivido grandes epidemias com expressivo aumento do número de casos e introdução de novos municípios com casos; atualmente apenas 10% dos municípios do Estado não apresentam casos autóctones. Observam-se picos epidêmicos aproximadamente a cada três anos bem como aumento do número de casos no ano subsequente (Figura3)(32).

Em 2011, o sorotipo DENV-4 foi isolado no Estado de São Paulo, em dois pacientes residentes em São José do Rio Preto e Paulo de Faria, ambos municípios localizados na região Noroeste do Estado (33,34).

No município de Campinas, em 1996 foi detectado o primeiro caso autóctone da doença (35). Desde o registro do primeiro caso autóctone, a doença vem sendo detectada anualmente na cidade, com a circulação do vírus da dengue dos tipos 1, 2 e 3 nestes anos de epidemia. Em 1998, ocorreu a primeira grande epidemia pelo sorotipo DENV-1, em 2002 pelo DENV-3 e pelo DENV-1 e DENV-2 em 2010, além de haver indícios de circulação do

DENV-2 nos anos de 1997 a 1999. O ano de 2007 foi marcado pela a maior epidemia de dengue na cidade, com incidência de 1.089,4 casos/100.000 habitantes (36).

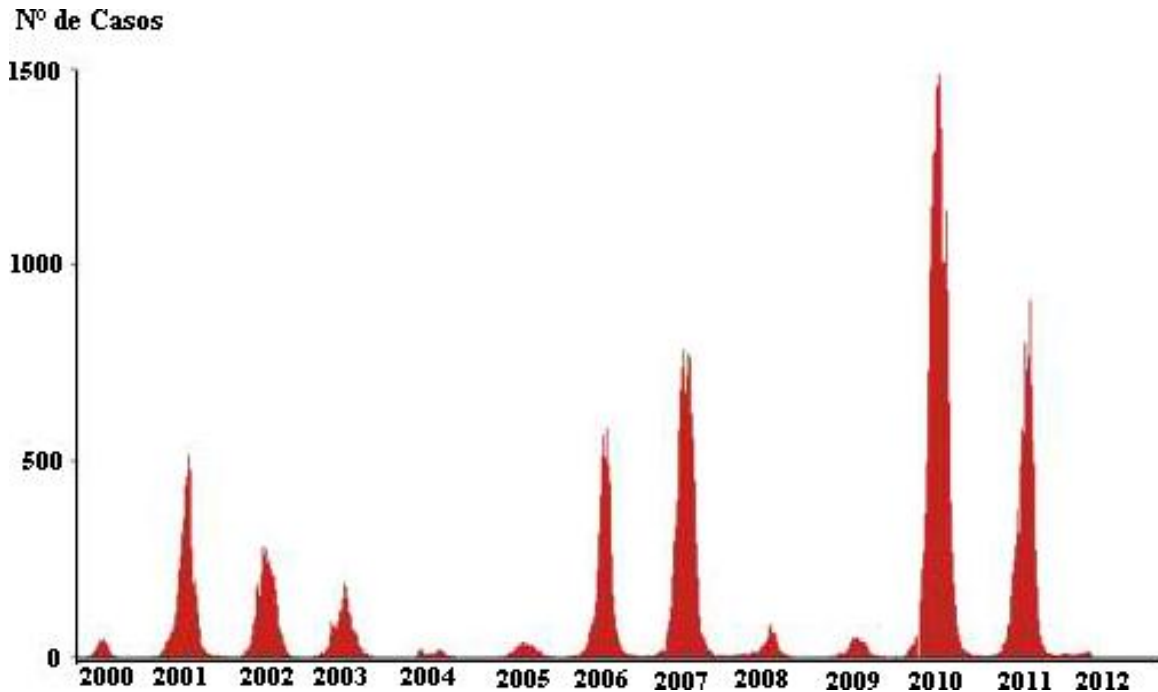


Figura 3. Distribuição dos casos autóctones de dengue por início dos sintomas, Estado de São Paulo.

Extraído de: Divisão de Zoonoses - CVE-SES/SP (32).

1.5. Espaço e Dengue

Sabe-se que as condições socioeconômicas e ambientais têm grande influência na produção, distribuição e propagação de doenças. O mundo moderno apresenta as condições favoráveis para a rápida expansão da dengue. A urbanização acelerada criou cidades com deficiências de abastecimento de água e de limpeza urbana, com intensa utilização de materiais não biodegradáveis, como recipientes descartáveis, de borracha, plástico e vidro, proporcionam condições ecológicas favoráveis à transmissão dos vírus da dengue pelo *Aedes aegypti* (1,3,15).

O aprofundamento no conhecimento das condições de vida e das prioridades comunitárias, implantação de políticas de saneamento eficaz, regularização do abastecimento de água e difusão de informações geograficamente localizadas, parecem ser caminhos que apontam, não só para reduzir vulnerabilidades à dengue e outras doenças, como também para uma melhoria da qualidade geral de vida da população(37).

Desde sua utilização inicial, o termo "endemia" associa-se a lugar. Assim, o estudo da ocorrência endemias, seus determinantes e seu processo de disseminação mistura-se ao estudo dos lugares e de seus habitantes, sendo impossível dissociá-los. No processo de compreensão de doenças como as vetoriais, o espaço deve necessariamente entrar como categoria de análise da epidemiologia, abordando transformações geradas pela ação humana sobre focos naturais (38).

A dimensão espacial é uma preocupação que se apresenta ao longo da história da epidemiologia e saúde pública, incluindo os registros que se referem ao pensamento de Hipócrates, que explorou a relação entre saúde e meio ambiente, com ênfase nas características físicas do ambiente e características culturais dos habitantes de um lugar. Destacando-se, nesse contexto, a notória e pioneira pesquisa de John Snow, nos estudos da cólera na Inglaterra em 1853 (39).

A evolução do conhecimento, sua especialização e segmentação em diferentes áreas, tem levado ao desenvolvimento de abordagens teóricas e metodológicas que têm contribuído para a compreensão da saúde a partir de perspectivas diferentes, como uma expressão da relação entre o espaço-sociedade (38).

A obra do geógrafo Milton Santos constitui uma das mais importantes referências para o estudo das relações entre espaço e o processo saúde-doença. Segundo, Santos, o espaço, nos primórdios da história humana, era compreendido como um mero conjunto de complexos naturais. Com o passar do tempo e a ação humana torna-se cada vez menos naturalizado e muito mais o resultado da produção histórica, tendendo a uma negação da “natureza natural”. O espaço é concebido como resultado da soma e da síntese, sempre refeita, da sociedade com a paisagem através da espacialidade. O espaço é a sociedade inserida na paisagem (38).

Assumindo-se o pressuposto de que a doença é considerada uma manifestação do indivíduo, as condições de vida são manifestações do lugar (38). O lugar, por sua vez, é pensado como um espaço organizado para análise e intervenções, no qual se deve buscar a identificação da situação de saúde com os seus determinantes culturais, ambientais e sociais (40).

Nesse sentido, a unidade espaço/população tem, então, a possibilidade de ser uma unidade em que operam os processos determinantes (condição de vida), expressam-se os problemas de saúde e desenvolvem-se as ações de saúde e bem-estar. O espaço geográfico é uma categoria de síntese que abriga a dimensão simbólica das relações sociais, na qual se expressam os fatores relacionados com o desenvolvimento das doenças e da sua distribuição entre os diversos grupos sociais. O espaço é construído socialmente e constitui, portanto, uma possibilidade de estratificar a população segundo a condição de vida (38,41).

A utilização das categorias condição de vida e espaço representam mediações passíveis de informar certas relações entre a saúde e a sociedade. A incorporação da categoria espaço nos estudos de saúde não só leva ao estabelecimento de diferenciações entre conjuntos de regiões, de acordo com características distintivas, como também permite a introdução da variável localização nos estudos. Este trabalho pressupõe o entendimento e discussão das diferenças entre regiões e a relação com o espaço no qual se inserem (42).

Diante disso, o estudo da localidade ganha importante destaque, uma vez que é neste nível que o processo de transmissão da doença se concretiza. Os estudos locais permitem a observação de variáveis e indicadores que em outros níveis de análise não seriam perceptíveis, uma vez que cada localidade possui uma historicidade própria, fruto de processos sociais e políticos singulares (43).

Para compreender como um contexto afeta a saúde de grupos populacionais através de processos de seleção, distribuição, interação, adaptação e outras respostas, é necessário medir efeitos em nível de grupo, uma vez que medidas em nível individual não podem dar conta destes processos (44).

Desta forma, as investigações sobre as condições particulares da ocorrência da dengue contribuem para o entendimento do papel dos grupos sociais, levando em conta os seus limites e possibilidades, na complexa dinâmica da cadeia de transmissão da doença,

podendo inclusive, acrescentar importantes elementos na discussão das estratégias de prevenção e controle mais adequadas (43).

1.6. Análise de dados espaciais em saúde

A análise espacial em saúde constitui o estudo quantitativo de fenômenos que são localizados geograficamente no espaço e podem ser realizadas através dos métodos de visualização, de análise exploratória ou de modelagem de dados espaciais (39,45). Essa abordagem de análise vem sendo empregada nos estudos epidemiológicos como um método que permite a integração de informações socioeconômicas, ambientais e demográficas, a fim de captar as desigualdades existentes, sem dissociá-las do espaço territorial (38,41,46).

O uso do geoprocessamento na área de saúde tem sido facilitado pelo amplo acesso a bases de dados epidemiológicos e pela disponibilização de ferramentas cartográficas e estatísticas computadorizadas. Estes avanços possibilitam a confecção rápida de mapas temáticos que podem contribuir para a formulação de hipóteses a respeito da distribuição espacial de agravos à saúde e sua relação com indicadores socioeconômicos (47). A definição de geoprocessamento encontra-se no conjunto de técnicas computacionais para coleta, tratamento, manipulação e apresentação de dados espaciais. Esse conjunto de técnicas aplicadas à saúde coletiva permite o mapeamento de agravos, a avaliação de riscos, o planejamento das ações em saúde e avaliação de redes de atenção, bem como as alternativas de intervenções (48).

A análise espacial utilizando geoprocessamento está muito mais acessível aos profissionais da saúde com a utilização de sistemas de informações nacionais e sistemas de informações geográficas (SIG). Os SIGs permitem a integração e combinação de dados cartográficos e seus atributos oriundos de diferentes fontes de dados (41). A incorporação desta ferramenta às técnicas epidemiológicas contribui para a análise de informações acerca dos eventos de saúde e condições socioeconômicas e ambientais das populações que ocupam diferentes espaços geográficos de uma cidade (38).

A inclusão, aplicação e o desenvolvimento do processamento e análise automatizada de bases de dados georreferenciadas tendem a aprimorar a compreensão do

espaço na produção e disseminação de doenças e agravos à saúde e constituem-se em mais uma ferramenta do processo de gestão dos riscos e do planejamento em saúde (46).

O estudo dos padrões de distribuição geográfica da incidência de doenças e suas relações com fatores socioambientais de risco, constituem-se no objeto do que hoje chama-se de Epidemiologia Geográfica ou Espacial. Apresenta quatro grandes áreas de interesse: mapeamento de doenças, estudos ecológicos, estudos de agregados (*cluster*) e avaliação e vigilância ambiental e tem se constituído em campo de aplicação de métodos e modelos de análise cada vez mais sofisticados na área da estatística (49).

Nos estudos espaciais, as unidades de análise definem a escala de observação dos fenômenos estudados, limitando as interações que podem ser captadas (41). Por sua vez, a escolha da escala de observação é limitada pela disponibilidade das unidades espaciais de referência nos sistemas de informações utilizados.

A aplicação das técnicas de análises espacial para compreensão do espaço na produção e disseminação de doenças e agravos à saúde sugere a aplicação de escalas ecológicas para abordar fatores que possam estar envolvidos nessas relações. Para entender como o contexto afeta a saúde de grupos populacionais por meio de distribuição, interação, adaptação e outras situações, faz-se necessário medir os efeitos no grupo, uma vez que as medidas individuais não conseguem esclarecer estes processos (44).

As técnicas de análise espacial envolvem as necessidades dos estudos ecológicos, que utilizam áreas geográficas como unidade usual de observação. O incremento das técnicas de análise espacial tem possibilitado a valorização dos estudos ecológicos em saúde, pois esse tipo de estudo esteve relegado à condição de abordagens descritivas sem poder analítico, ao longo dos tempos (39,49).

No caso particular da análise espacial da ocorrência de doenças, onde se dispõe de dados referidos a uma área geográfica, tem-se presenciado, principalmente na última década, o uso crescente dos chamados Modelos Hierárquicos Bayesianos com o auxílio de métodos de simulação (50).

Essas técnicas partem do pressuposto básico de que taxas e indicadores epidemiológicos de uma doença são realizações de um processo estocástico que se procura identificar (50). Ao se trabalhar com processos intramunicipais, o uso de pequenas áreas

para a agregação de dados referentes a eventos raros gera grande variabilidade aleatória, que não reflete os reais diferenciais de risco (39).

Assim, para identificar padrões de distribuição espacial e de risco para doenças a partir de pequenas unidades territoriais de análise (como setores censitários, p.ex.), supõe-se que estes riscos estejam conectados por uma relação com o que acontece na vizinhança de cada unidade espacial. Em decorrência dessa relação entre as unidades espaciais, é necessário realizar uma ponderação das taxas calculadas pelos valores dos vizinhos, de forma a permitir a estabilização estatística desses indicadores, além de uma visão mais próxima do processo gerador da realização observada (39).

Para a análise espacial, conceitos como dependência espacial e autocorrelação espacial são fundamentais. Entende-se por dependência espacial o fato de que a maior parte das ocorrências naturais ou sociais apresentam entre si uma relação que depende da distância. A estrutura de dependência espacial é evidenciada a partir da autocorrelação espacial, que é uma medida da estrutura de dependência espacial (39,51,52).

1.7. Modelos Inferenciais Bayesianos

Para a Inferência Clássica, toda informação sobre o parâmetro a ser estimado está contida na amostra. O modelo probabilístico é estabelecido sobre os dados amostrais, considerando-se que os parâmetros são desconhecidos. Este modelo é o que se chama Função de Verossimilhança, uma função escrita em termos dos dados e dos parâmetros. O método de estimação conhecido como Método da Máxima Verossimilhança consiste em encontrar os valores dos parâmetros que maximizam a função de verossimilhança estabelecida com aqueles dados amostrais. Isto é, observada a amostra, quais os valores dos parâmetros tornam mais prováveis os dados coletados? Estes valores são chamados de estimativas de máxima verossimilhança. A estas estimativas, podem ser associados intervalos de confiança, expressando a incerteza embutida no processo de estimação dos parâmetros (53).

Na Inferência Bayesiana, nem toda informação sobre o parâmetro vem da amostra. Parte dela está contida na amostra e é expressa através da função de verossimilhança. Outra parte reside num conhecimento a priori sobre o parâmetro (antes de

se conhecer dados) e pode ser expressa através de uma distribuição de probabilidade, a chamada distribuição a priori do parâmetro, ou simplesmente, priori do parâmetro. O peso de cada informação, amostral ou da priori, é regulado através de quão informativa é a distribuição a priori atribuída ao parâmetro. As duas partes da informação são combinadas com o auxílio do Teorema de Bayes, resultando na distribuição a *posteriori* do parâmetro. Esta distribuição é utilizada nas inferências sobre o parâmetro. A partir da distribuição do parâmetro, conhecido os dados (a *posteriori*), é possível calcular probabilidades, percentis, média, entre outras propriedades (53).

A distribuição a *priori* dos parâmetros de interesse também depende de outros parâmetros, os hiperparâmetros, denominados assim para serem diferenciados dos parâmetros principais. A estes hiperparâmetros, podem ser atribuídos valores ou outras distribuições a priori, as chamadas *hiperprioris*. Este processo continua até um nível em que valores são atribuídos aos hiperparâmetros. A referência a este processo como Modelagem Hierárquica é bastante comum.

Se existe nenhum ou pouco conhecimento a *priori* sobre o parâmetro, costuma-se expressar esta ignorância através da atribuição de uma distribuição a priori conhecida como priori vaga, na qual a variância é grande, deixando um amplo domínio para a variação dos valores do parâmetro (53).

2. OBJETIVOS

2.1. Objetivo Geral

Analisar a distribuição espacial da incidência de dengue na cidade de Campinas, no ano de 2007 e verificar a sua relação com condições socioambientais.

2.2. Objetivos Específicos

1. Estudar a distribuição espacial da incidência de dengue segundo os grupos etários até 14 anos e acima de 14 anos de idade, por áreas de cobertura dos centros de saúde.
2. Analisar por pequenas áreas a distribuição espacial dos casos de dengue e sua relação com condições socioambientais e avaliar a adequação de diferentes modelos para dados de contagem na presença de excesso de zeros.

3. MATERIAIS E MÉTODOS

3.1. Desenho do estudo

Foi realizado um estudo ecológico espacial com os casos de dengue agregados, utilizando-se os setores censitários e centros de saúde como unidades de análise.

3.2. Área de estudo

O município de Campinas está situado na região centro-oeste do Estado de São Paulo (47°04'40'' Longitude Oeste e 22°53'20'' Latitude Sul), numa altitude média de 640 metros acima do nível do mar. O município apresenta área total de 795,7 km², com 49% de área urbana e população segundo o Censo 2010 igual a 1.080.113 habitantes, dos quais 20,3% são menores de 15 anos e 98,3% da população total estão estabelecidas na área urbana.

A gestão da saúde no município apresenta-se subdividida em cinco Distritos de Saúde (DS), subdivididos em Áreas de cobertura dos centros de saúde (ACCS). Em 2007, nesta configuração contabiliza-se 47 ACCS. Cada ACCS caracteriza-se com delimitação territorial definida pela facilidade de acesso da população, dimensionadas para prestar assistência e ações de saúde pública a aproximadamente 20.000 habitantes. Periodicamente corrige-se a subdivisão em ACCS conforme o crescimento populacional da cidade, em que em 2012 o município apresentava-se com 61 ACCS.

3.3. Dados

Casos Notificados de Dengue

Foram analisados todos os casos de dengue autóctone confirmados laboratorialmente registrados entre 01 de janeiro e 31 de dezembro de 2007, por ser o ano de maior incidência da doença na cidade. Os dados foram obtidos do Sistema de Informação de Agravos de Notificação (SINAN), fornecido pela Coordenadoria de Vigilância e Saúde Ambiental da Secretaria Municipal de Saúde de Campinas (Covisa/SMS). As variáveis coletadas do SINAN incluídas para análise espacial foram: endereço de residência (logradouro e numeração) e idade. Ressalta-se que de nenhuma forma os nomes

dos indivíduos que se apresentaram como casos foram identificados, analisando-se somente a contagem dos casos.

Após a revisão e correção dos dados notificados, os endereços dos casos foram georreferenciados com o software ArcMap 9.2 (Environmental Systems Research Institute (ESRI), Inc., Redlands, CA), segundo os parâmetros: Elipsóide Internacionald Hayford, Projeção Universal de Mercator, Datum Horizontal Córrego Alegre, MG (maiores detalhes ver Malavasi, 2011). Posteriormente, os casos foram agregados segundo duas diferentes bases cartográficas digitalizadas de Campinas, a saber: a malha de polígonos das ACCS, para o ano de 2007, fornecida pela Secretaria Municipal de Saúde de Campinas e segundo a malha de setores censitários, com a divisão do IBGE para o ano de 2000.

Para agregação dos casos de dengue por ACCS, criaram-se três distintos bancos de dados: 1) o total de casos foi agregado segundo doze quadrissemanas epidemiológicas do período; 2) agregados segundo o grupo etário menor ou igual a 14 anos e 3) agregados segundo grupo etário com idade superior a 14 anos. Para os setores censitários criou-se uma única base de dados. No processo de geocodificação de pontos para polígonos utilizou-se o pacote *Maptools* do *software* R 2.15.0 (R Development Core Team, 2012).

Índice de carência socioambiental

No processo de modelagem espacial, considerou-se por variável independente a variável categórica índice de carência socioambiental (ICSA). Esta é uma covariável sintética que permite diferenciar as distintas áreas geográficas em estudo (ACCS e setores censitários), em função de suas condições socioambientais. O ICSA foi elaborado a partir de variáveis do Censo 2000, já utilizadas em estudos anteriores (54,55) e que demonstraram alguma relação com a incidência de dengue. Foram utilizadas as seguintes variáveis do Censo 2000 na elaboração da covariável ICSA:

- **Renda domiciliar:** Proporção de pessoas responsáveis por domicílios com renda de até três salários-mínimos.
- **Escolaridade:** Proporção de pessoas responsáveis por domicílios com até três anos de estudo

- **Esgotamento sanitário:** Proporção de domicílios com esgoto inadequado (Fossa rudimentar+Vala+Rio+Outro escoadouro)
- **Limpeza urbana:** Proporção de domicílios com destinação do lixo inadequada (Queimado+Enterrado+Terreno baldio ou logradouro, Jogado em rio, lago ou mar e Outro destino).

Para tanto, foi empregada a técnica estatística de análise de componentes principais (ACP), a partir do procedimento PRINCOMP do software SAS v. 9.2 (SAS Institute, Cary, NC). Em seguida, considerando-se o primeiro componente principal obtido com a ACP e o procedimento FASTCLUS do SAS, foi realizada a análise de agrupamento (cluster analysis), pelo método não-hierárquico *k-means*, com o objetivo de agregar as ACCS em estratos (*clusters*) com a máxima homogeneidade interna e máxima heterogeneidade entre os estratos (56).

Ao final, a variável do ICSA foi incorporada a dois bancos de dados com as contagens dos casos, relacionando-os por ACCS. Para possibilitar estimar as taxas de incidência de dengue, em cada um desses dois bancos, também foi incorporada a contagem populacional de cada ACCS, estimada pela Coordenadoria de Informação e Informática da SMS Campinas, a partir de dados do Censo.

3.4. Estimação do risco relativo para dados agregados

Os modelos lineares generalizados (GLM) são uma extensão dos modelos lineares clássicos, sendo indicados quando as variáveis em estudo não têm aderência a distribuição normal, principalmente pelo fato de serem processos de contagem. Estes modelos compõem um grupo de distribuições de probabilidades conhecido como família exponencial de distribuições e englobam diversas funções aditivas como, a regressão linear, de Poisson, logística, log-linear, entre outras (57). Apresentam uma estrutura de linearidade na média e consistem basicamente em três componentes:

- Uma distribuição de probabilidade da variável resposta na família exponencial;
- Um preditor linear $\eta = \sum_{i=1}^k z_i \beta_i$;

c) Uma função de ligação g tal que $\eta = g(\mu)$.

Pode-se demonstrar que as distribuições Normal, Binomial, Bernoulli, Poisson, Exponencial e Gama pertencem a família exponencial (57). Assim, modela-se o valor esperado da variável resposta como função das covariáveis, através de uma função utilizada como ligação.

$$\eta_i = \alpha + \mathbf{x}'_i \boldsymbol{\beta} \quad i = 1, \dots, N \quad (1)$$

Na equação (1), α denota o intercepto (valor médio do risco) e o vetor $\boldsymbol{\beta}$ representa os efeitos lineares das covariáveis \mathbf{x} . No mapeamento de doenças, seja a região de estudo dividida em N áreas contíguas e Y_i representa a contagem dos casos de dengue (variável resposta) indexada a área $i = 1, \dots, N$. Assume-se que Y_i segue uma distribuição da família exponencial. A média μ_i é relacionada ao preditor linear η_i , através de uma função de ligação conhecida $g(\cdot)$, em que $g(\mu_i) = \eta_i$ e $\mu_i = \exp(\eta_i)$, envolvendo os efeitos de varias covariáveis de forma aditiva (57).

Na modelagem espacial de doenças, se a doença é rara, o modelo de Poisson $Y_i | \theta_i \sim \text{Poisson}(\mu_i = E_i \theta_i)$ é o modelo de referência para analisar as Y_i contagens de doença (39,45,58), em que E_i e θ_i são, respectivamente, o número esperado de casos e o risco relativo desconhecido da doença, na i -ésima área ($i=1, \dots, N$). O modelo padrão de regressão Poisson é definido por:

$$\Pr(Y_i = y_i | \mu_i) = \frac{\mu_i^{y_i} \exp(-\mu_i)}{y_i!} \quad \text{para } y_i \geq 0 \quad (2)$$

$$\log(\mu_i) = \mathbf{X}'_i \boldsymbol{\beta}$$

$$\mu_i = \exp(\mathbf{X}'_i \boldsymbol{\beta}) \quad (3)$$

$$\mathbf{X}'_i \boldsymbol{\beta} = \beta_0 + X_{ICSA} \boldsymbol{\beta}_{ICSA}$$

em que \mathbf{X}_{ICSA} é uma a matriz de planejamento para duas categorias da covariável ICSA (a categoria “menor” é assumida como referência), β_0 é o intercepto do modelo a ser estimado, $\boldsymbol{\beta}_{ICSA}$ é um vetor de parâmetros desconhecido que descreve o efeito fixo das condições socioambientais sobre o risco de dengue.

O estimador de máxima verossimilhança de μ_i , denotado por $\hat{\mu}_i = SIR_i = Y_i/\mu_i$, é a taxa de incidência padronizada, uma taxa bruta, que apresenta instabilidade frente a pequenos valores observados ou para populações pequenas ou ambos (51,52,59). O modelo de Poisson pressupõe igualdade da média e da variância de Y_i . Contudo, em razão da presença de superdispersão dos dados, dificilmente esta situação é verificada (60). E ainda, é pouco realista a suposição de homogeneidade dentro de cada área, em que os riscos individuais são esperados iguais dentro das áreas. Assim, destacamos a seguir algumas extensões deste modelo básico que possibilitam modelar esta variação extra-Poisson. Primeiramente, apresentam-se os modelos BN e ZI e posteriormente discuti-se a inclusão de efeitos aleatórios segundo o enfoque bayesiano.

Modelo Binomial Negativo

O modelo BN é o principal modelo em substituição à regressão de Poisson. É formulado a partir de uma mistura das distribuições de Poisson e Gama. Tem por finalidade capturar a superdispersão do modelo, modelando explicitamente os eventos correlacionados via uma variável latente com distribuição Gama (60).

Clayton e Kaldor (59) foram os primeiros a adotarem um modelo de efeito aleatório (ou mistura) que assume uma função densidade de probabilidade paramétrica para a distribuição do risco relativo entre áreas. Supõe-se que o risco relativo é independente e identicamente com distribuição de mistura Gama conjugada a um parâmetro:

$$Y_i|\mu_i \sim \text{Poisson}(v_i\mu_i) \quad (4)$$

$$v_i \sim \text{Gama}(\delta, \delta) \quad (5)$$

No modelo BN, δ é um parâmetro de escala e representa a superdispersão dos dados. Assim, a distribuição marginal de Y é uma distribuição BN, expressa por $Y_i|\mu_i, \delta \sim BN(\mu_i, \delta)$, com a seguinte função de probabilidade:

$$P(Y_i = y_i|\mu_i, \delta) = \frac{\Gamma(y_i + \delta)}{\Gamma(y_i + 1)\Gamma(\delta)} \cdot \left(\frac{\mu_i}{\mu_i + \delta}\right)^{y_i} \cdot \left(\frac{\delta}{\mu_i + \delta}\right)^\delta, \quad \delta > 0 \quad (6)$$

em que $\Gamma(\cdot)$ representa uma função Gama(\cdot). A média e a variância são:

$$E(y_i|\mu_i, \delta) = \mu_i,$$

$$Var(y_i|\mu_i, \delta) = \mu_i + \mu_i^2/\delta \quad (7)$$

em que $\mu_i = \exp(\mathbf{X}'_i\beta)$, conforme descrito na equação (3).

Verifica-se que a variância é igual à média multiplicada por um fator de dispersão $\varphi_i = (1 + \mu_i/\delta)$. Assim, o modelo BN possibilita que a variância seja superior à média, acomodando o problema de superdispersão e observa-se que a distribuição BN converge a uma distribuição de Poisson quando δ tende a infinito.

Modelo Inflacionado de Zeros (ZI)

Os modelos ZI surgem quando a variável resposta é modelada como uma mistura de uma distribuição de Bernoulli para a presença/ausência de zeros e outra distribuição para os valores positivos (Poisson ou BN). O excesso de zeros é ajustado assumindo que os zeros surgem a partir de dois distintos processos (61). O primeiro processo ocorre com probabilidade ω_i e produz apenas verdadeiros zeros (ou zeros estruturais), o segundo, denominado falsos zeros (ou zeros amostrais) ocorrem com probabilidade $(1 - \omega_i) \Pr(Y_i = 0)$, em que $\Pr(Y_i = 0)$ é uma distribuição de Poisson ou BN. A probabilidade total de zeros é a combinação das probabilidades dos dois processos, ou seja, $\Pr(Y_i = 0) = \omega_i + (1 - \omega_i) \Pr(Y_i = 0)$.

Modelo de Poisson Inflacionado de Zeros (ZIP)

O modelo BN descrito acima é capaz de ajustar a superdispersão presente nos dados, entretanto, este modelo não é apropriado para modelar a superdispersão decorrente de um número excessivo de contagens com valor zero nos dados (60). Para modelar este excesso de zeros, a principal solução é estimarmos um modelo que misture a Poisson com uma distribuição que leva em conta o excesso de zeros. Nesse contexto, podemos aplicar o modelo de regressão de Poisson inflacionado de zeros (ZIP). Desta forma, tem-se a hipótese que, com probabilidade ϖ_i , a variável resposta assume o valor zero, e com

probabilidade $(1 - \varpi_i)$ assume o valor de uma variável aleatória com distribuição Poisson, de média μ_i , conforme já descrito anteriormente.

O modelo ZIP (60) considera que os zeros verdadeiros ocorrem com probabilidade ϖ_i e os falsos zeros ocorrem com probabilidade $(1 - \varpi_i)$. O modelo ZIP é expresso da seguinte forma:

$$P(Y_i = y_i | \varpi_i) = \begin{cases} \varpi_i + (1 - \varpi_i)\exp(-\lambda_i), & y_i = 0 \\ (1 - \varpi_i) \frac{\lambda_i^{y_i} \exp(-\lambda_i)}{y_i!}, & y_i > 0. \end{cases} \quad (8)$$

O parâmetro ϖ_i tem a restrição $0 \leq \varpi_i < 1$ e $\lambda_i > 0$. A média e a variância de Y_i são, respectivamente, $E[Y_i] = (1 - \varpi_i)\lambda_i$ e $Var[Y_i] = (1 - \varpi_i)\lambda_i + \varpi_i(1 - \varpi_i)\lambda_i^2$.

Podemos observar que a variância da mistura é maior que a média da distribuição. Quanto maior a probabilidade do excesso de zeros, maior a variância da variável. À medida que p se aproxima de zero, a variância se aproxima de μ_i , ou seja, voltamos a lidar somente com uma distribuição Poisson padrão.

Seus parâmetros $\boldsymbol{\varpi} = (\varpi_1, \dots, \varpi_n)'$ e $\boldsymbol{\lambda} = (\lambda_1, \dots, \lambda_n)'$ são modelados usualmente pelas funções de ligação canônica logarítmica, $\log(\lambda_i) = \mu + \mathbf{B}\boldsymbol{\beta}$, e logística, $\text{logito}(\varpi_i) = \alpha + \mathbf{G}\boldsymbol{\gamma}$, em que μ e α são dois interceptos, \mathbf{B} e \mathbf{G} são duas matrizes das variáveis independentes (com o conjunto destas variáveis sendo iguais ou não), e $\boldsymbol{\beta}$ e $\boldsymbol{\gamma}$ são dois vetores de parâmetros desconhecidos dos modelos de regressão.

Se o processo de contagem não segue o modelo de Poisson, então, pode-se usar o modelo binomial negativo inflacionado de zeros (ZINB), considerando o processo de contagem como uma distribuição binomial negativa.

Modelo Binomial Negativo Inflacionado de Zeros (ZINB)

A distribuição ZINB (62,63) surge como uma mistura da distribuição Binomial Negativa e uma distribuição que leve em conta o excesso de zeros. O modelo ZINB pode ser escrito como segue:

$$P(Y = y_i) \begin{cases} \varpi_i + (1 - \varpi_i) \left(\frac{\delta}{\delta + \lambda_i} \right)^\delta, & y_i=0 \\ (1 - \varpi_i) \frac{\Gamma(\delta + y_i)}{\Gamma(\delta)\Gamma(y_i + 1)} \left(\frac{\delta}{\delta + \lambda_i} \right)^\delta \left(\frac{\lambda_i}{\delta + \lambda_i} \right)^{y_i}, & y_i>0 \end{cases} \quad (9)$$

A média e a variância de Y_i é:

$$E[Y_i] = (1 - \varpi_i)\lambda_i \text{ e}$$

$$Var[Y_i] = (1 - \varpi_i)\lambda_i + (1 + \lambda_i(\varpi_i + 1/\delta)). \quad (10)$$

A distribuição ZINB aproxima-se da ZIP quando $\delta \rightarrow 0$ e aproxima-se da binomial negativa quando $\varpi_i \rightarrow 0$. Se ambos $1/\delta$ e ϖ_i convergem para zero, então a distribuição ZINB é reduzida a distribuição de Poisson padrão. Assim, conforme o modelo ZIP, as funções de ligação canônica são logarítmica, $\log(\lambda_i) = \mu + \mathbf{B}\boldsymbol{\beta}$, e a logística, $\text{logito}(\varpi_i) = \alpha + \mathbf{G}\boldsymbol{\gamma}$, em que a interpretação destes parâmetros é análoga ao modelo ZIP.

Nesta tese, nos modelos ZIP e ZINB, descritos acima, os zeros verdadeiros representam a situação em que não houve registros de casos de dengue em um dado setor, mesmo com a presença de uma população humana suscetível e com a presença do vetor e/ou patógeno no ambiente. Isto pode ser devido à resistência imunológica da população ou porque a transmissão não ocorreu. Ou seja, a situação era favorável para a incidência da doença, mas ela não ocorreu. Os falsos zeros representam a situação em que a população não se encontrava suscetível a doença, em virtude da ausência do vetor e/ou patógeno no ambiente ou a situação em que realmente houve ocorrência de casos de dengue em um dado setor, mas nenhum caso foi registrado, possivelmente por subnotificação (64).

3.5. Modelos Hierárquicos Bayesianos

No contexto de inferência bayesiana, todos os parâmetros do modelo θ são considerados como variáveis aleatórias, em que a distribuição que representa o conhecimento sobre θ (a quantidade desconhecida) é dada por $p(\theta)$, denominada de distribuição a priori e expressa o conhecimento ou incerteza sobre θ , antes da obtenção dos dados amostrais \mathbf{y} (53).

Seja θ um vetor de parâmetros de interesse e que representa um objeto de estudo (por exemplo, o risco relativo de uma doença) e seja y o vetor dos dados observados para a variável dependente e x um conjunto de variáveis explicativas. A modelagem bayesiana inicia com a especificação de um modelo probabilístico completo $p(\theta/y)$. O conhecimento a priori sobre θ é resumido na densidade $p(\theta)$ e a informação de θ contida nos dados de x é a verossimilhança $p(y/\theta)$. Estas duas fontes de informações (*a priori* e a verossimilhança) estão relacionadas por probabilidades condicionais, dada pelo Teorema de Bayes, em que o conhecimento atualizado está contido em uma densidade conhecida como distribuição *a posteriori* (53). Pelo Teorema de Bayes temos:

$$p(\theta/y) = \frac{p(y/\theta)p(\theta)}{p(y)}, \quad (11)$$

em que $p(y) = \int_{\Theta} p(\theta)p(y/\theta)d\theta$ é a verossimilhança marginal e Θ é o espaço de parâmetros. Uma forma equivalente de (11) omite o fator $p(y)$, que não depende de θ . Assim a distribuição *a posteriori* pode ser reescrita como:

$$p(\theta/y) \propto p(y/\theta)p(\theta) \quad (12)$$

Estas expressões resumem a essência da inferência bayesiana. A tarefa fundamental de qualquer aplicação específica é desenvolver o modelo $p(\theta/y)$ e realizar os cálculos necessários para representar o conhecimento *a posteriori* $p(\theta/y)$ de forma apropriada.

As inferências intervalares – na abordagem bayesiana denominadas intervalos de credibilidade (IC) ou intervalos de probabilidade *a posteriori* – derivam da *posteriori* e são usualmente obtidas através de delimitações nos quantis. O IC assume que o verdadeiro valor do parâmetro está contido nesse intervalo – que representa um intervalo de probabilidade – especificado com probabilidade pré-determinada, considerando-se o parâmetro supostamente aleatório (53).

3.6. Efeitos aleatórios

A aplicação dos modelos de contagens, descritos anteriormente, tem como premissa a independência das observações. Entretanto, as estimativas obtidas com esta abordagem não leva em consideração a possível dependência espacial entre as áreas geográficas próximas, em que áreas vizinhas, com características ambientais (físicas e sócio-econômicas) semelhantes, tendem a apresentar correlação entre os riscos.

Uma das alternativas apresentadas na literatura para superar esta dificuldade é a aplicação de métodos hierárquicos bayesianos (50). Os métodos bayesianos espaciais de mapeamento de doenças supõem comumente que as áreas são condicionalmente independentes, dados os parâmetros do modelo, e incorporam subsequentemente a dependência como parte de uma distribuição dos parâmetros. A informação *a priori* permite a inserção do conhecimento a respeito do efeito espacial no modelo, através da imposição de determinada estrutura de covariância. Consequentemente, é através da modelagem da distribuição *a priori* que podemos introduzir a dependência espacial entre os riscos.

O modelo proposto por Besag *et al.* (denotado por modelo BYM) (65), permite modelar a dependência espacial a partir da incorporação de duas componentes de efeitos aleatórios: 1) espacialmente estruturado ($f_i(s)$), que introduz estrutura espacial ao modelo, permite a cada área compartilhar a informação de áreas vizinhas (subjacentes), em que o valor esperado deste efeito aleatório condicionado em cada área em relação as outras áreas, é obtido a partir da média do efeito nas áreas vizinhas, e 2) efeito de heterogeneidade não estruturado $f_i(u)$, que é independente para cada área. Assim, o modelo descrito na equação, considerando-se o efeito da covariável ICSPA e dos efeitos aleatórios tem a seguinte formulação:

$$\eta_i = \alpha + \sum_{k=1}^{p-1} \beta_k ICSPA_k + f_i(s) + f_i(u) \quad (13)$$

$$\mu_i = \exp \left(\alpha + \sum_{k=1}^{p-1} \beta_k ICSPA_k + f_i(s) + f_i(u) \right), \quad k = 1, \dots, p-1 \quad (p = 3) \quad (14)$$

em que α é o intercepto e representa o risco médio global, β_k é o vetor de efeitos fixos das $ICSa_k$ categorias da variável de condições socioambientais, em que as estimativas do risco relativo consideram as comparações com a p -ésima categoria de condições socioambientais, $f_i(u)$ corresponde a heterogeneidade não estruturada devido a diferença entre os riscos nas diferentes áreas e $f_i(s)$ são efeitos espacialmente estruturados devido a autocorrelação espacial entre as áreas.

Segundo a abordagem bayesiana hierárquica (50), distribuições *a priori* para o efeito espacialmente estruturado são representadas por um campo aleatório Markoviano Gaussiano (CAMG), cuja estrutura é definida usando-se um modelo autoregressivo condicional (CAR) Gaussiano intrínseco (66). Esta *a priori* indica a relação espacial de vizinhança entre as áreas. Considera-se que duas áreas são vizinhas se elas compartilham alguma relação de adjacência. A intensidade do processo espacial $f_s(s_i)$ sobre uma vizinhança é definida por.

$$f_i^{(s)} \mid \mathbf{f}_{-i}^{(s)}, \boldsymbol{\tau}^{(s)} \sim N\left(\frac{1}{n_i} \sum_{j \in \partial_i} f_j^{(s)}, \frac{1}{n_i \boldsymbol{\tau}^{(s)}} \right) \quad (15)$$

em que ∂_i representa o conjunto de áreas (setores censitários) vizinhas à área i e n_i corresponde ao número de áreas vizinhas à área i . Os valores para $f_j^{(s)}$ são 1 se duas áreas são vizinhas e 0 caso contrário. A especificação da distribuição CAR depende de um único parâmetro $\boldsymbol{\tau}^{(s)} = 1/\sigma_s^2$, em que $\boldsymbol{\tau}^{(s)}$ e σ_s^2 representam a precisão do processo e a variabilidade condicional estruturada espacialmente, respectivamente. A distribuição CAR é imprópria e como *prioris* impróprias podem resultar em *posterioris* impróprias, o modelo da equação (1) pode apresentar problemas de identificabilidade dos parâmetros (50). Para contornar este problema, impõe-se uma restrição de soma zero para o efeito $f_i^{(s)}$, ou seja, $\sum_{i=1}^n f_i^{(s)} = 0$. No nível dos hiperparâmetros, os parâmetros de precisão desconhecidos $\boldsymbol{\tau}^{(s)}$ e $\boldsymbol{\tau}^{(u)}$, são especificados de acordo com *hiperprioris* não-informativas. Frequentemente, atribui-se para estas *hiperprioris* uma distribuição Gama, tal como: $\boldsymbol{\tau}^{(s)}, \boldsymbol{\tau}^{(u)} \sim G(0,5; 0,0005)$ (67,68). Essa especificação permite obter maior variabilidade nas estimativas dos efeitos aleatórios, possibilitando que a verossimilhança dos dados domine a informação

contida na *priori*, portanto, esta *priori* terá pouco efeito sobre a inferência da distribuição *a posteriori* dos riscos relativos. No nível dos parâmetros, α , β_s e γ_s foram especificados por *prioris* com distribuições Gaussianas não informativas.

3.7. Método INLA

Para a obtenção da distribuição *a posteriori*, utilizou-se o método INLA (*Integrated Nested Laplace Approximations*) (69,70), em que, aproximações para as distribuições marginais *a posteriori* são obtidas em três etapas. Na primeira, aproximam-se as distribuições marginais *a posteriori* dos hiperparâmetros θ , ou seja, os parâmetros da distribuição *a priori* do campo latente \mathbf{x} , a partir da relação:

$$\pi(\theta|\mathbf{y}) \propto \frac{\pi(\mathbf{x}, \theta, \mathbf{y})}{\pi(\mathbf{x}|\theta, \mathbf{y})}, \quad (16)$$

em que θ representa os hiperparâmetros, \mathbf{y} os dados e \mathbf{x} o campo latente descrito acima.

$\pi(\theta|\mathbf{y})$ pode ser aproximado por:

$$\tilde{\pi}(\theta|\mathbf{y}) \propto \frac{\pi(\mathbf{x}, \theta, \mathbf{y})}{\tilde{\pi}_G(\mathbf{x}|\theta, \mathbf{y})} \Bigg|_{\mathbf{x}=\mathbf{x}'(\theta)}, \quad (17)$$

em que o denominador $\tilde{\pi}_G(\mathbf{x}|\theta, \mathbf{y})$ é uma aproximação Gaussiana para a distribuição condicional completa de \mathbf{x} , a moda $\mathbf{x}^*(\theta)$ é avaliada para um dado valor dos hiperparâmetros θ .

No segundo passo é calculada uma aproximação da distribuição *a posteriori* do campo latente \mathbf{x} , dadas as \mathbf{y} variáveis respostas observadas e os hiperparâmetros θ . Este pode ser estimado por

$$\tilde{\pi}(x_i|\theta, \mathbf{y}) \propto \frac{\pi(\mathbf{x}, \theta|\mathbf{y})}{\tilde{\pi}_{GG}(\mathbf{x}_{-i}|x_i, \theta, \mathbf{y})} \Bigg|_{\mathbf{x}_{-i}=\mathbf{x}_{-i}^*(x_i, \theta)}, \quad (18)$$

em que $\mathbf{x}_{-i}^*(x_i, \theta)$ corresponde a configuração modal de $\pi(\mathbf{x}_{-i}|x_i, \theta, \mathbf{y})$ e $\tilde{\pi}_{GG}(\mathbf{x}_{-i}|x_i, \theta, \mathbf{y})$ é a aproximação Gaussiana para $\pi(\mathbf{x}_{-i}|x_i, \theta, \mathbf{y})$, que difere da distribuição condicional obtida a partir de $\tilde{\pi}_G(\mathbf{x}|\theta, \mathbf{y})$.

O terceiro passo é a realização de integração numérica com respeito a $\boldsymbol{\theta}$, que consiste em encontrar as distribuições a *posteriori* dos componentes do campo latente \mathbf{x} dada as variáveis resposta observada \mathbf{y} variáveis, como a soma da seguinte forma:

$$\tilde{\pi}(x_i|\mathbf{y}) = \sum_k \tilde{\pi}(x_i|\boldsymbol{\theta}_k, \mathbf{y})\tilde{\pi}(\boldsymbol{\theta}_k|\mathbf{y}) \Delta_k , \quad (19)$$

Na equação (19), Δ_k é o peso atribuído a área correspondente ao ponto de integração $\boldsymbol{\theta}_k$, $\tilde{\pi}(\boldsymbol{\theta}_k|\mathbf{y})$ é a distribuição marginal posterior aproximada dos hiperparâmetros $\boldsymbol{\theta}$ no ponto de integração k e $\tilde{\pi}(x_i|\boldsymbol{\theta}_k, \mathbf{y})$ é a distribuição marginal a *posteriori* aproximada do campo latente no ponto de integração k.

3.8. Comparação dos modelos bayesianos

O critério de informação da *deviance* (DIC) é definido a partir da distribuição a *posteriori* da estatística *deviance*: $D(\boldsymbol{\theta}) = -2 \log p(\mathbf{y}|\boldsymbol{\theta}) + 2 \log f(\mathbf{y})$. Em que, $p(\mathbf{y}|\boldsymbol{\theta})$ é a função de verossimilhança para o vetor de dados observados \mathbf{y} , dado o vetor de parâmetros $\boldsymbol{\theta}$ e $f(\mathbf{y})$ é uma função padronizada dos dados (que não tem nenhum impacto sobre a seleção do modelo). Nesta abordagem, a seleção do modelo é resumida pela esperança a *posteriori* da *deviance* $D = E_{\boldsymbol{\theta}|\mathbf{y}}[D]$, a complexidade do modelo é capturada pelo número efetivo de parâmetros p_D , definido por: $p_D = E_{\boldsymbol{\theta}|\mathbf{y}}[D] - D(E_{\boldsymbol{\theta}|\mathbf{y}}[\boldsymbol{\theta}]) = 2\bar{D} - D(\bar{\boldsymbol{\theta}})$. A medida p_D , frequentemente é menor que o número total de parâmetros do modelo. O DIC, então é definido por:

$$DIC = \bar{D} + p_D = 2\bar{D} + p_D, \quad (20)$$

em que o menor valor do DIC indica o modelo com melhor ajuste.

A escolha do modelo mais adequado também pode ser realizada utilizando-se o critério da Ordenada Preditiva Condicional (CPO – *conditional predictive ordinate*) (71). Este critério é uma medida preditiva e pode ser obtido facilmente com o método INLA. O valor da CPO obtido por $CPO_{ijk} = \pi(y_{ijk,obs}|\mathbf{y}_{-(i,j,k),obs})$ é definido como a densidade preditiva de validação cruzada $\pi(y_{ijk}|\mathbf{y}_{-(i,j,k)})$ na observação $y_{ijk,obs}$, em

que $\mathbf{y}_{-(i,j,k)}$ denota os dados sem ijk -ésima observação. Para os dados desta tese utilizamos o logaritmo da CPO definido por:

$$\overline{LCPO} = -\frac{1}{n} \sum_{i,j,k} \log(CPO_{ijk}) \quad (21)$$

O menor valor de \overline{LCPO} indica o melhor modelo.

4. RESULTADOS

4.1. Artigo 1

Aceito para publicação na revista “Cadernos de Saúde Pública”

**DISTRIBUIÇÃO ESPACIAL DA INCIDÊNCIA DE DENGUE E
CONDIÇÕES SOCIOAMBIENTAIS EM CAMPINAS, ESTADO DE SÃO
PAULO, BRASIL, 2007**

SPATIAL DISTRIBUTION OF DENGUE INCIDENCE AND SOCIAL AND
ENVIRONMENTAL CONDITIONS IN CAMPINAS, STATE OF SÃO
PAULO, BRAZIL, 2007

José Vilton Costa¹, Maria Rita Donalizio¹, Lician Vaz de Arruda Silveira²

¹Faculdade de Ciências Médicas - UNICAMP. Departamento de Saúde Coletiva.

²Instituto de Biociências – UNESP/Botucatu

RESUMO

O objetivo deste estudo foi analisar a distribuição espacial do risco de dengue e a sua relação com condições socioambientais. Trata-se de um estudo ecológico da contagem dos casos de dengue autóctone, no município de Campinas, Estado de São Paulo, ano de 2007, agregados em 47 áreas de cobertura dos Centros de Saúde do município. Modelos espaciais de mapeamento de doenças foram construídos utilizando-se modelos hierárquicos bayesianos, através do método de Integração Aproximada Aninhada de Laplace (INLA). As análises foram estratificadas segundo os grupos etários até 14 anos e acima de 14 anos. Os resultados indicam que a distribuição espacial do risco de dengue não está associada a condições socioambientais para o grupo etário até 14 anos. No grupo etário acima de 14 anos, o risco relativo de dengue aumenta significativamente conforme aumenta o nível de carência socioambiental. O mapeamento dos estratos de carência socioambiental e dos casos de dengue mostrou-se uma ferramenta útil na análise dos dados dos sistemas de vigilância da dengue.

Palavras chaves: Dengue; Fatores Socioambientais; Análise Espacial; Modelos Bayesianos.

ABSTRACT

This study aimed to analyze the spatial distribution of the risk of dengue fever and its relation to socio-environmental conditions. It is an ecological study of the count autochthonous dengue cases in the municipality of Campinas, São Paulo, 2007, aggregated into 47 coverage areas of the health centers in the city. Spatial models for disease mapping were constructed using a Bayesian approach based on the Integrated Nested Laplace Approximation (INLA). The analyzes were stratified according to age groups 0 to 14 years and above 14 years. The results indicate that the spatial distribution of dengue risk is not associated with social-environmental deprivation for the age 0 to 14 years. In the age group above 14 years, the relative risk of dengue increases significantly with increasing the level of socio-environmental conditions. The mapping of socio-environmental deprivation strata and the dengue cases has proved a useful tool in the analysis of data from surveillance systems for dengue.

Key words: Dengue; Socio-environmental Factors; Spatial Analysis; Bayesian models.

Introdução

A dengue é uma doença viral de transmissão vetorial causada por um dos quatro sorotipos do vírus dengue (DENV-1 a DENV-4), pertencentes ao gênero *Flavivirus*, da família *Flaviviridae*. Atualmente a dengue é considerada um dos maiores problemas de saúde pública com cerca de 100 milhões de casos anualmente no mundo, sendo que mais de 2,5 bilhões de pessoas vivem em área sob risco de infecção. É endêmica de regiões tropicais como o sudeste asiático, sul do Pacífico, África Oriental, Caribe e América Latina¹. No Brasil, a região Sudeste é a que registra o maior número de casos de dengue por ano, sendo que as demais regiões, por ordem de incidência de casos de dengue são Nordeste, Centro-Oeste, Sul e Norte^{1,2}.

A crescente circulação dos vários sorotipos, de diferentes genótipos do vírus como também a (re)introdução de novos sorotipos virais da dengue no Brasil tem resultados na “endêmicação” e hiperendemicidade da doença com mudanças no seu perfil epidemiológico. Esta tendência tem sido relacionada com o aumento da incidência em menores de 14 anos e da proporção de casos graves³.

No município de Campinas, estado de São Paulo, o primeiro registro de um caso de dengue autóctone foi em 1996. Desde então, a doença vem sendo detectada anualmente, tendo ocorrido três epidemias com maior intensidade de transmissão nos anos de 1998, 2002 e 2007. Neste último ano foram contabilizados 11.519 casos da doença, sendo detectada a circulação de três sorotipos diferentes do vírus DENV-1, DENV-2 e DENV-3. A série histórica de casos, desde 1998, mostra que o período de maior incidência da doença ocorre nos meses de março e abril em Campinas⁴, como também observado em outros municípios⁵.

As razões para a re-emergência e expansão da dengue nas últimas décadas no Brasil são complexas. Alguns autores apontam a associação entre a incidência da doença e as condições climáticas (aumento da temperatura, pluviosidade e umidade do ar)^{6,7}. Por outro lado, o acelerado crescimento populacional, intensos fluxos migratórios para as periferias urbanas sem infraestrutura e a deterioração das condições de vida e saúde são

fatores contribuintes para o aumento da infestação do principal vetor da dengue, o *Aedes aegypti*^{1,6} e a consequente dispersão da doença^{1,8}.

Atualmente, sem tratamento específico e ainda sem uma vacina eficaz disponível contra a dengue, uma das formas de prevenção existente consiste no controle do mosquito transmissor⁹. Assim, compreender as dimensões espaciais da transmissão da doença em diferentes contextos epidemiológicos aumenta o entendimento sobre a dinâmica de transmissão do vírus, ajudando o planejamento de medidas de controle preventivas e emergenciais¹⁰.

As diferentes dimensões da transmissão da dengue, resultantes da interação entre as populações humanas, os vetores e a circulação dos vírus exigem abordagens de variadas disciplinas e áreas de conhecimento. Dentre elas a análise espacial permite avaliar a distribuição de eventos de saúde em diferentes níveis de agregação geográfica¹⁰⁻¹². Têm sido utilizadas para avaliar a distribuição da dengue no espaço urbano, na identificação de regiões de sobrerisco e na busca de fatores potencialmente associados com o risco em diferentes regiões¹³⁻¹⁴.

O objetivo deste estudo é analisar a distribuição espacial da incidência de dengue nas áreas de cobertura dos centros de saúde em Campinas e avaliar possíveis relações com indicadores socioambientais.

Materiais e Métodos

Local e desenho de estudo

Trata-se de um estudo ecológico sobre a distribuição espacial dos casos de dengue autóctone, em moradores do município de Campinas, tendo como unidade espacial de análise as Áreas de Cobertura dos Centros de Saúde (ACCS) do município de Campinas. Adotou-se esta unidade de análise em decorrência da sua importância como território delimitado para o planejamento e gestão dos serviços de saúde, além de *locus* onde se observa a interação dos serviços com a população organizada (conselhos locais de saúde).

A cidade está situada na região centro-oeste do Estado de São Paulo (47°04'40'' Longitude Oeste e 22°53'20'' Latitude Sul), numa altitude média de 640

metros acima do nível do mar. O município apresenta área total de 795,7 km², com 49% de área urbana e população de 1.080.113 habitantes, dos quais 20,3% são menores de 15 anos e 98,3% da população total estão estabelecidas na área urbana.

A gestão da saúde no município está subdividida em cinco Distritos de Saúde (DS) subdivididos em ACCS. Cada uma das 47 áreas possui delimitação territorial definida pela facilidade de acesso da população, dimensionadas para prestar assistência e ações de saúde pública a aproximadamente 20.000 habitantes.

Os dados de dengue

Foram analisados todos os casos de dengue confirmados laboratorialmente registrados entre 01 de janeiro e 31 de dezembro de 2007, por ser o ano de maior incidência da doença na cidade. Os dados foram obtidos do Sistema de Informação de Agravos de Notificação (SINAN), fornecido pela Coordenadoria de Vigilância e Saúde Ambiental da Secretaria Municipal de Saúde de Campinas (Covisa/SMS). As variáveis coletadas do SINAN incluídas para análise espacial foram: endereço de residência (logradouro e numeração) e idade.

Após a revisão e correção dos dados notificados¹⁵, os endereços dos casos foram georreferenciados com o software ArcMap 9.2 (ENSRI), segundo os parâmetros: Elipsóide Internacionald Hayford, Projeção Universal de Mercator, Datum Horizontal Córrego Alegre, MG. Após a obtenção das coordenadas de latitude e longitude, foi realizada geocodificação dos casos em base cartográfica digitalizada de Campinas, contendo a malha das ACCS fornecida pela Secretaria Municipal de Saúde de Campinas, utilizando o pacote Maptools do *software* R 2.15.0 (*R Development Core Team*, 2012).

O banco de dados resultante nesse processo de geocodificação foi subdividido em dois, de acordo com o grupo etário dos casos: até 14 anos e maiores de 14 anos de idade. Em uma etapa final, em cada um desses dois bancos de dados, os casos foram agregados por ACCS.

Estratos de carência socioambiental

Para diferenciar as ACCS em função de suas condições socioambientais, elaborou-se uma classificação das 47 ACCS a partir da construção de um índice de carência socioambiental (ICSA) para cada ACCS. Esta classificação foi capaz de sintetizar variáveis do Censo 2000, algumas já utilizadas em estudos anteriores^{14,16,17}. Foram utilizados os seguintes indicadores: proporção de pessoas responsáveis por domicílios com renda até três salários-mínimos; proporção de pessoas responsáveis por domicílios com até três anos de estudo; proporção de domicílios com esgotamento sanitário inadequado e proporção de domicílios com destinação do lixo inadequada.

Para tanto, foi utilizada a técnica estatística de componentes principais (ACP), a partir do procedimento PRINCOMP do software SAS v. 9.2 (SAS Institute, Cary, NC). Em seguida, considerando-se o primeiro componente principal obtido com a ACP e o procedimento FASTCLUS do SAS, foi realizada a análise de agrupamento (*cluster analysis*), pelo método não-hierárquico *k-means*, com o objetivo de agregar as ACCS em estratos (*clusters*) com a máxima homogeneidade interna e máxima heterogeneidade entre os estratos¹⁸.

Ao final, a variável do ICSA foi incorporada a dois bancos de dados com as contagens dos casos, relacionando-os por ACCS. Para possibilitar estimar as taxas de incidência de dengue, em cada um desses dois bancos, também foi incorporada a contagem populacional de cada ACCS, estimada pela Coordenadoria de Informação e Informática da SMS Campinas, a partir de dados do Censo.

Análise espacial

Para cada $ACCS_i$ ($i=1,\dots,47$) foram calculadas razões padronizadas de incidência (Standardized Incidence Ratio – SIR), segundo o método de padronização indireta. Consideramos como referência a taxa global de incidência em Campinas.

A modelagem hierárquica bayesiana foi empregada para modelar a distribuição espacial do risco relativo de dengue nas ACCS e para investigar a associação com o ICSA.

No mapeamento de doenças com dados agregados, uma abordagem clássica na análise do risco relativo é a aplicação do modelo proposto por Besag, York e Mollié (modelo BYM)¹⁹. Este modelo assume que as contagens y_i são observações condicionalmente independentes com distribuição de Poisson, com a seguinte formulação:

$$y_i | \eta_i \sim P(\exp(\eta_i)) \quad i = 1, \dots, 47 \quad (1)$$

em que y_i são os casos observados de dengue e η_i é o preditor aditivo estruturado²⁰ e denota o logaritmo do risco relativo para a i -ésima ACCS. O preditor η_i é expandido para acomodar a covariável do ICSA, assumindo-se a seguinte forma:

$$\eta_i = \alpha + \beta_1 \text{ICSA}(\text{moderada})_i + \beta_2 \text{ICSA}(\text{maior})_i + f_i^{(u)} + f_i^{(s)} \quad (2)$$

em que α é o intercepto para o log do risco relativo de todas as ACCS, β_1 e β_2 são os parâmetro de efeitos fixos da variável ICSA. Os termos $f_i^{(u)}$ e $f_i^{(s)}$ representam efeitos aleatórios independentes que modelam a heterogeneidade não estruturada espacialmente e estruturada espacialmente, respectivamente.

Desde que estamos usando a abordagem bayesiana, *prioris* foram especificadas para os parâmetros $\alpha, \beta_1, \beta_2, f_i^{(u)}$ e $f_i^{(s)}$. Assumimos para os parâmetros α, β_1 e β_2 uma distribuição *a priori* não informativa Gaussiana, com média igual a zero e precisão (inverso da variância) igual a 0,00005²¹. O parâmetro de precisão foi adotado o menor possível para refletir a nossa falta de conhecimento prévio sobre a força de associação entre o risco de dengue e o ICSA.

Para os componentes aleatórios, os elementos do vetor de efeitos não estruturados $\mathbf{f}(u)$ são independentes e atribui-se uma distribuição *a priori* especificada como $\mathbf{f}(u) \sim N(0, \tau_u = 1/\sigma_u^2)$, em que σ_u^2 é a variabilidade não-estruturada. O vetor de efeito estruturado espacialmente, $\mathbf{f}(s)$, é ajustado por um campo aleatório Markoviano Gaussiano (CAMG)²⁰. O CAMG tem sua estrutura determinada por um modelo autoregressivo condicional (CAR) Gaussiano intrínseco:

$$f_i^{(s)} | \mathbf{f}_{-i}^{(s)}, \tau^{(s)} \sim N\left(\frac{1}{n_i} \sum_{j \in \partial_i} f_j^{(s)}, \frac{1}{n_i \tau^{(s)}}\right) \quad (3)$$

em que ∂_i representa o conjunto de ACCS vizinhas à ACCS i e n_i corresponde ao número de ACCS vizinhas à ACCS i . A especificação da distribuição CAR depende de um único parâmetro $\tau^{(s)} = 1/\sigma_s^2$, em que $\tau^{(s)}$ e σ_s^2 representam a precisão e a variabilidade condicional estruturada espacialmente, respectivamente. A distribuição CAR é imprópria e como *prioris* impróprias podem resultar em *posteriors* impróprias, o modelo da equação (2) pode apresentar problemas de identificabilidade dos parâmetros. Para contornar este problema, impõe-se uma restrição de soma zero para o efeito $f_i^{(s)}$, ou seja, $\sum_{i=1}^{47} f_i^{(s)} = 0$.

Para finalizar a especificação dos parâmetros do modelo, os dois hiperparâmetros de precisão desconhecidos $\tau^{(s)}$ e $\tau^{(u)}$, são especificados de acordo com *hiperprioris* não-informativas com distribuição gama $G(0,5; 0,0005)$, conforme recomendação padrão da literatura^{22,23}. Essa especificação permite obter maior variabilidade nas estimativas dos efeitos aleatórios, possibilitando que a verossimilhança dos dados domine a informação contida na *priori*, portanto, esta *priori* terá pouco efeito sobre a inferência da distribuição *a posteriori* dos riscos relativos.

Todos os parâmetros foram estimados usando inferência bayesiana, a partir de aproximações analíticas segundo a metodologia INLA (*Integrated Nested Laplace Approximations*)²⁴. Este método foi desenvolvido para permitir a realização de inferência bayesiana aproximada, dentro da classe de modelos gaussianos latentes, possibilitando aproximar marginais *a posteriori* de interesse. O seu desenvolvimento surge como uma alternativa aos métodos de Monte Carlo via cadeias de Markov (MCMC), que têm sua forma de inferência via simulações. O método INLA, por sua vez, realiza aproximações de forma direta às distribuições *a posteriori* de interesse. Com isso, é possível realizar inferência com um custo computacional consideravelmente inferior além de não ser preciso lidar com problemas de convergência.

No processo de modelagem do risco relativo de dengue, as análises foram estratificadas por dois grupos etários: até 14 anos e acima de 14 anos de idade, e foram construindo os seguintes modelos por estrato populacional:

$$M1: Y \sim f_i^{(u)}$$

$$M2: Y \sim f_i^{(s)}$$

$$M3: Y \sim f_i^{(u)} + f_i^{(s)}$$

$$M4: Y \sim f_i^{(u)} + f_i^{(s)} + \text{ICSA}$$

As faixas etárias foram assim estratificadas devido à crescente incidência e gravidade da dengue em menores de 14 anos³.

Os modelos foram comparados e selecionados usando o critério de informação da *deviance* (DIC). Os menores valores do DIC indicam o melhor ajuste do modelo²². Além disso, para cada ACCS foi calculada e mapeada a probabilidade *a posteriori* do risco relativo (RR) de dengue ser maior do que um ($\text{Pr}[\text{RR} > 1 | Y]$), obtendo-se assim, uma medida de evidência (estatística) de excesso de risco em cada área. Estimativas de probabilidades *a posteriori* superiores a 0,8 representam evidências de um excesso de risco. Estudos de simulação²⁵ demonstram que a classificação das áreas em “excesso de risco”, com probabilidade *a posteriori* acima do limiar de 0,8 apresenta boa capacidade preditiva para detectar regiões com verdadeiros RR ao redor de 1,5 a 2,0 quando as contagens esperadas em cada região são em torno de 20 ou para detectar RR próximo de 3, quando as contagens esperadas são pequenas. A modelagem bayesiana foi realizada com o *software* R 2.15.0 (R Development Core Team, 2012), utilizando o pacote R-INLA (disponível em: <http://www.r-inla-org>).

Resultados

No ano de 2007 foram notificados pelo Sistema de Vigilância Epidemiológica em Campinas 11.519 casos de dengue, resultando em taxa de incidência de 1089,4 casos por 100 mil habitantes. Desse total de casos notificados, o processo de georreferenciamento conseguiu identificar 8982 (78%) endereços de residência sobre o eixo de ruas do município. A não identificação do restante dos casos deveu-se à discrepância entre o endereço e a base cartográfica.

O perfil etário dos casos mostrou que 15,8% tinha 14 anos ou menos, com uma taxa de 612,2 casos por 100 mil habitantes. Nos maiores de 14 anos, a incidência foi 1027,3 casos por 100 mil habitantes.

A criação do ICSA resultou na estratificação das ACCS em 3 grandes grupos de carência socioambiental. Das 47 ACCS, 17 (36%) foram classificadas no estrato de menor carência socioambiental; 21 (45%) foram classificadas no estrato moderado e 9 (19%) no estrato de maior carência.

Contrastando-se as taxas de incidência pelo ICSA e por grupo etário, observa-se que a taxa de incidência de dengue cresce num gradiente de intensidade do menor para o maior nível de carência socioambiental, para ambos os grupos etários e para o total dos casos. (Tabela 1).

Tabela 1. População, número de casos e incidência de dengue segundo o índice de carência socioambiental (ICSA) e grupo etário. Campinas-SP, 2007.

Grupo etário	ICSA	População	Total de Casos	Taxa de Incidência (por 100 mil/hab.)
≤ 14 anos	Menor carência	63460	254	400,3
	Moderada	104246	666	638,9
	Maior carência	64409	501	777,8
>14 anos	Menor carência	291643	1943	666,2
	Moderada	306500	3681	1201,0
	Maior carência	137899	1937	1404,7
Total	Menor carência	355103	2197	618,7
	Moderada	410746	4347	1058,3
	Maior carência	202308	2438	1205,1

A distribuição espacial do índice de carência socioambiental, por ACCS (Figura 1), mostra a heterogeneidade das condições de vida da população em Campinas. Nota-se um gradiente crescente, em termos de carência socioambiental, do centro à periferia.

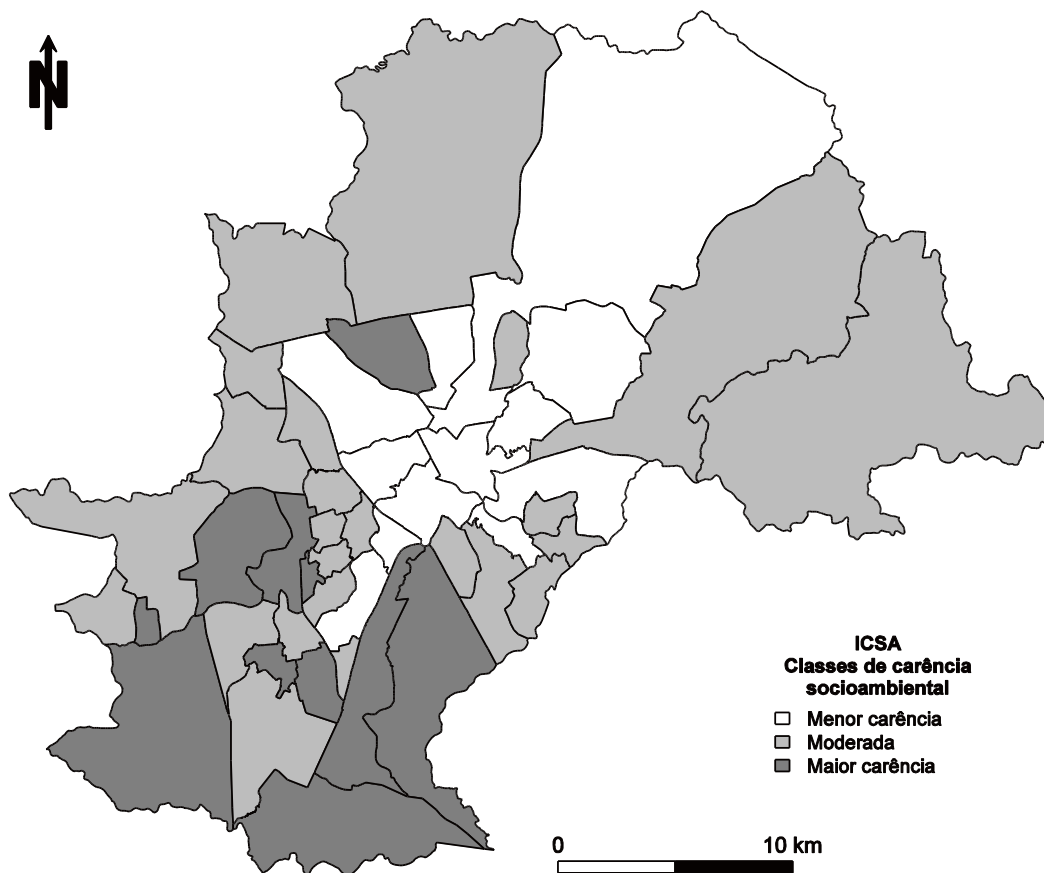


Figura 1. Distribuição espacial do índice de carência socioambiental (ICSA), segundo classes de carência, por área de cobertura dos centros de saúde em Campinas-SP, 2007.

A Tabela 2 apresenta os modelos avaliados pelo modelo BYM e respectivos valores do critério de informação da *deviance* (DIC). Verifica-se que para os casos de dengue no grupo etário ≤ 14 anos, os quatro modelos apresentaram valores do DIC muito próximos. Considerando-se que valores menores do DIC indicam melhor ajuste do modelo, verificou-se que a inclusão do efeito estruturado espacialmente melhorou o ajuste dos modelos, por outro lado, a inclusão da covariável ICSA no modelo M4 não melhorou o ajuste. O modelo M3, ajustado para os efeitos aleatórios não estruturado $f(u)$ e estruturado espacialmente $f(s)$ apresentou menor valor do DIC, ou seja, foi o modelo com melhor ajuste. Assim, deste ponto em diante, ao nos reportamos à análise do risco relativo de dengue no grupo etário ≤ 14 anos, estaremos utilizando apenas o modelo M3.

Tabela 2. Modelos ajustados e critério de Informação da *Deviance* (DIC). Campinas-SP, 2007.

Grupo etário	Modelos ajustados	Critério de informação da <i>Deviance</i> (DIC)
≤ 14 anos	M1: f(u)	416,35
	M2: f(s)	414,76
	M3: f(u)+f(s)	413,71
	M4: f(u)+f(s)+ICSA	413,84
> 14 anos	M5: f(u)	416,72
	M6: f(s)	415,41
	M7: f(u)+f(s)	413,93
	M8: f(u)+f(s)+ICSA	412,89

f(u): efeito aleatório não estruturado.

f(s): efeito aleatório estruturado espacialmente.

ICSA: índice de carência socioambiental.

Os modelos ajustados para os casos ocorridos no grupo etário > 14 anos também apresentaram os valores do DIC bem próximos. O modelo M8, ajustado para o efeito da covariável ICSA e para os efeitos aleatórios f(u) e f(s), foi aquele que apresentou melhor ajuste. Os dados da Tabela 3 representam as estimativas a posteriori do risco relativo de dengue segundo o efeito fixo do ICSA. As estimativas do risco relativo, nesse grupo etário de > 14 anos, aumentam conforme passamos da categoria de menor para maior carência socioambiental. O risco relativo e respectivos intervalos de credibilidade (IC), nas ACCS com carência moderada e com maior carência foram, respectivamente, 1,77 (IC95%: 1,33; 2,79) e 2,24 (IC95%: 1,27; 3,95) vezes maior em relação ao risco nas ACCS com menor carência socioambiental.

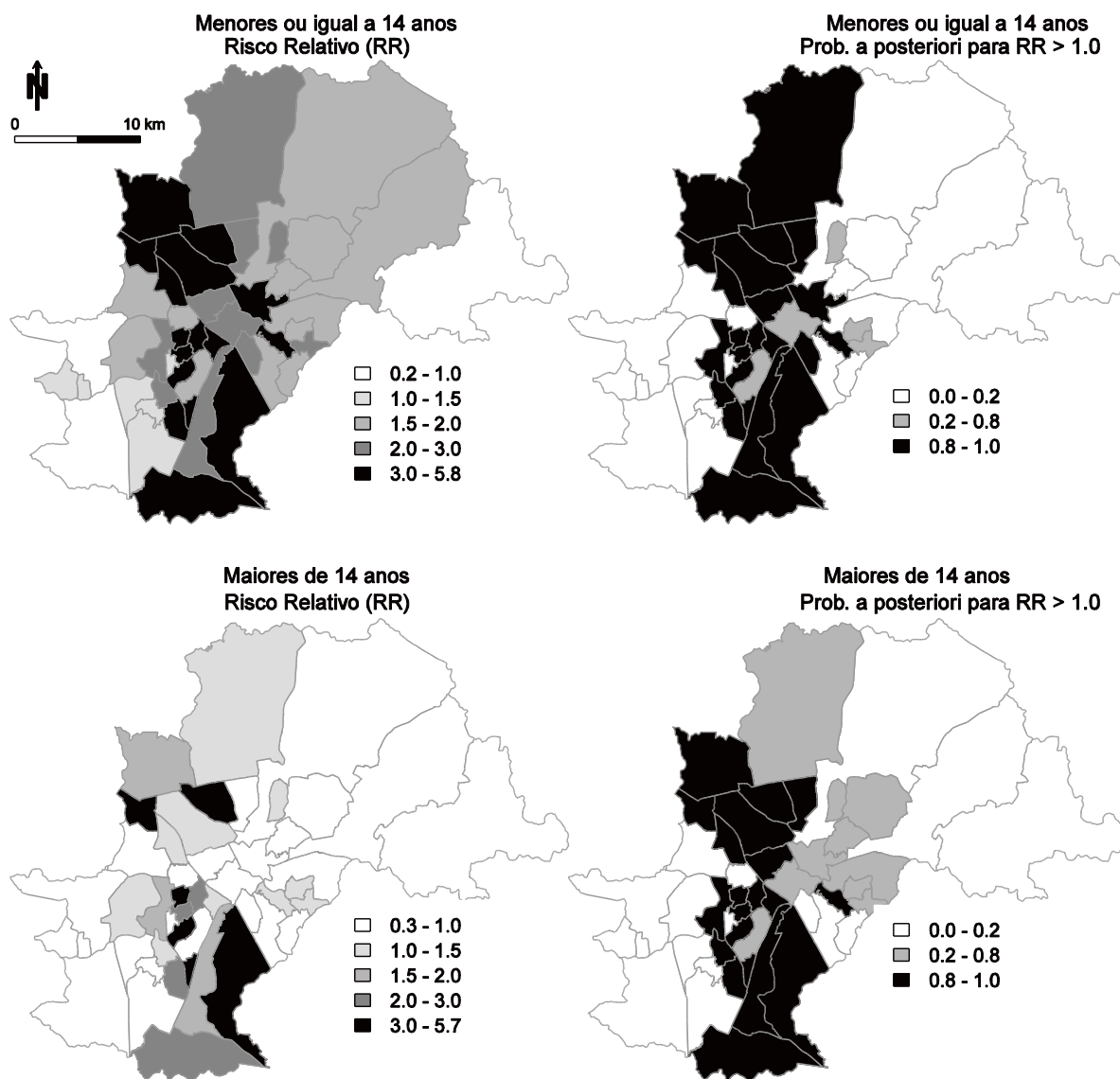


Figura 2. Distribuição espacial do risco relativo (RR) e probabilidade (Prob.) a posteriori do RR ser maior do que 1, para os grupos etários ≤ 14 anos ($f(u)+f(s)$) e > 14 anos de idade ($ICSA+f(u)+f(s)$). Campinas-SP, 2007.

A Figura 2 apresenta as estimativas suavizadas do risco relativo de dengue e probabilidades *a posteriori* do risco ser maior do que 1, para os grupos etários ≤ 14 anos e > 14 anos, obtidas segundo o modelo BYM. As estimativas do risco relativo variaram de 0,2 a 5,8 e de 0,3 a 5,7 nos grupos etários ≤ 14 anos e > 14 anos, respectivamente. Ainda

sobre a distribuição do risco relativo, observa-se para ambos os grupos etários, que os maiores riscos foram identificados em ACCS localizadas nas regiões sul e norte.

O padrão da distribuição espacial do risco de dengue, por grupo etário (Figura 2), mostrou-se semelhante em relação à identificação das ACCS com excesso de risco ($\Pr[RR>1 | Y] > 0,8$). Estas ACCS identificadas apresentam excesso de risco para a população como um todo, indiferente do seu grupo etário. Observou-se excesso de risco em 17 ACCS para o grupo etário ≤ 14 anos e em 23 para o grupo etário >14 anos. São áreas que devem ter uma atenção especial no controle da dengue no município de Campinas.

Discussão

Os resultados deste estudo revelaram diferenças na distribuição espacial do risco de dengue para pequenas áreas do município de Campinas e que indicadores socioambientais estão associados com a incidência da dengue na cidade. A distribuição espacial das taxas de incidência de dengue, segundo o ICSA, aponta maior risco da doença em áreas com maior carência socioambiental em relação a áreas com menor.

As ACCS da região sul, em especial, apresentam desde 1999, altas taxas de incidência da doença na cidade²⁶. Embora na epidemia de 2007 tenham sido registradas as maiores taxas de incidência em Campinas, as precárias condições socioambientais na região sul e a maior circulação viral em períodos anteriores, com diminuição de susceptíveis podem estar associadas ao maior risco da doença em menores de 14 anos.

O estudo da circulação viral em anos anteriores poderia esclarecer algumas áreas pouco afetadas pela epidemia em 2007.

As precárias condições de vida em bairros da periferia da cidade podem ser fatores moduladores da maior incidência de dengue. As ACCS com maior risco relativo e com excesso de risco na porção norte da cidade são áreas também com precárias condições de infraestrutura urbana, localizadas na divisa com o município de Sumaré, áreas conurbadas com registro de altas taxas de incidência de dengue nas últimas décadas. A região sul da cidade é de urbanização recente, com áreas de invasão, grande concentração populacional, sendo que sua população apresenta precários indicadores socioeconômicos e

de infraestrutura urbana, como saneamento básico, coleta de lixo, intermitência no abastecimento de água etc²⁶.

Em regiões com falta de infraestrutura urbana e pobreza, o impacto das medidas de controle de vetores pode estar limitado, sem resultados diretos ou indiretos na eliminação de focos de *Aedes aegypti*, altamente adaptado aos centros urbanos^{1,27,28}.

Em um estudo caso-controle espacial realizado na região sul de Campinas, observou-se que fatores relacionados ao armazenamento de água, à frequência de coleta de lixo e à forma de esgotamento sanitário foram fatores associados à incidência de dengue²⁶. Embora a proliferação do vetor não esteja diretamente relacionada à maior cobertura de esgotamento sanitário, este fator indiretamente revela condições precárias de moradia e acesso a equipamentos e serviços urbanos²⁸. Estes resultados confirmam investigações nacionais^{16,17,29,30} e internacionais que relacionam a incidência de dengue a condições socioambientais²¹⁻³⁴. A maior circulação de vírus nestas áreas resulta em maiores taxas de incidência em crianças, residentes em regiões com acesso aos serviços também precários.

Por outro lado, Mondini e Chiaravalloti-Neto¹⁶ (2007) não encontram correlação espacial entre incidência de dengue e fatores socioeconômicos em município do estado de São Paulo. Resultados semelhantes também foram apontados no Rio de Janeiro³⁵, Nova Iguaçu, RJ³⁶ e Queensland, na Austrália³⁷.

Os resultados contraditórios encontrados nos diferentes estudos demonstram a complexa relação entre fatores socioambientais e o risco de infecção da doença, cuja transmissão também sofre interferência de fatores como a imunidade populacional aos vários sorotipos circulantes nesta epidemia e em anos anteriores³⁷. Alguns autores encontraram evidência da associação entre risco de dengue e condições socioambientais após a reintrodução de novo sorotipo do vírus, sendo que depois da dispersão do vírus as incidências ocorrem de forma geral em todos os extratos socioeconômicos³⁸. Vale destacar que este estudo analisou áreas no município onde a doença já havia circulado em anos anteriores, certamente interferindo nas taxas de incidência em 2007.

Os resultados discrepantes entre condições de vida e incidência também encontram explicação no próprio processo de produção da doença, para a qual as condições no microambiente doméstico podem ser mais relevantes do que as características contextuais

do ambiente em nível macro social³⁹. Por outro lado, observam-se entre os estudos, diferentes níveis de agregação espacial de dados epidemiológicos, além da utilização de indicadores de condições de vida e saúde com variadas capacidades de discriminação e aproximação com exposições contextuais de interesse na transmissão da dengue³⁹. A escolha da ACCS como a unidade territorial de análise facilitou a identificação de áreas delimitadas para a atuação de equipes específicas de controle de vetores e vigilância epidemiológica do município.

Embora existam outros fatores para explicar a incidência da doença, os dados sugerem que os fatores socioambientais fazem parte desse fenômeno.

A ausência de uma vacina eficaz para a prevenção da dengue desloca os esforços para o controle do vetor transmissor². As maiores taxas de incidência observadas na região sul de Campinas, repetindo o que já tinha ocorrido em epidemias nos anos de 1999 e 2002, apontam as limitações das intervenções na manutenção de situações de risco de proliferação de vetores e transmissão da doença.

Por um lado, as ações de combate ao vetor, como as inspeções domiciliares, a eliminação e tratamento de criadouros em conjunto com a educação em saúde e a mobilização social podem resultar na interrupção do ciclo de transmissão da dengue. Por outro lado, investimentos no saneamento básico, abastecimento de água, limpeza urbana, moradia e acesso à educação também têm impacto no controle da doença⁴⁰.

Este estudo destaca-se pelo uso de técnicas de geoprocessamento e análise espacial, com a integração entre os dados socioambientais, disponíveis no Censo Populacional 2000 (IBGE), e a base de dados dos casos de dengue, registradas no SINAN. Estas técnicas permitem levantar hipóteses sobre possíveis fatores de risco da doença, identificando áreas delimitadas de maior risco de forma oportuna para a intervenção⁴¹. A unidade espacial adotada neste estudo, ACCS, identifica áreas de risco de dengue que coincidem com as áreas utilizadas na organização de ações de prevenção e controle de doenças e de atuação dos serviços de saúde no município.

Destaca-se também a aplicação de modelos espaciais bayesianos hierárquicos sensíveis para avaliar a incidência de dengue e estabelecer relações com indicadores socioambientais no espaço intramunicipal do município de Campinas. A modelagem

estatística adotada permitiu controlar a correlação espacial entre ACCS vizinhas e comparar estatisticamente o risco de dengue em áreas geográficas com maior carência em relação a outras com menor carência socioambiental.

Finalmente, vale destacar que por se tratar de um estudo ecológico há dificuldade na generalização dos resultados e inferência causal, pois está sujeito a viés ecológico ou problemas da unidade de área modificável (MAUP)⁴², isto é, diferentes formas de agregação espacial dos casos podem influenciar os efeitos da relação entre fatores de risco e doença. Portanto, sugere-se estudar os efeitos das condições socioambientais considerando-se diferentes níveis de agregação geográfica e também a dimensão temporal. Além disso, destaca-se a carência de informações sobre a circulação dos diferentes sorotipos nas áreas de estudo dificultando a interpretação da distribuição espacial da doença e sua correlação com a intensidade da transmissão.

Conclusões

O mapeamento dos estratos de carência socioambiental e das taxas de incidência de dengue mostrou-se uma ferramenta útil na análise dos dados dos sistemas de vigilância da doença. As estimativas do risco relativo revelaram-se associadas às áreas de maior carência socioambiental. O índice de carência socioambiental foi significativo, mas a interpretação do seu efeito necessita ser combinado a outros tipos de informação sobre a ecologia do vetor, sorotipos circulantes e imunidade da população.

Referências

1. Brasil. Ministério da Saúde.
http://portal.saude.gov.br/portal/arquivos/pdf/casos_de_dengue_classica_brasil_1990_2011.pdf, acessado em 20/Jun/2012.
2. Gubler DJ: Dengue and dengue hemorrhagic fever: its history and resurgence as a global public health problem. In: Gubler DJ, Kuno G (Eds). Dengue and Dengue Hemorrhagic Fever. CAB International Press; 1997:1-22.
3. Teixeira MG, Costa MC, Barreto F, Barreto ML: Dengue: twenty-five years since reemergence in Brazil. *Cad Saúde Pública* 2009, 25 Suppl 1:S7-18.

4. Prefeitura Municipal de Campinas. Secretaria Municipal De Saúde. Coordenadoria de Vigilância em Saúde. Informe. Dengue: Situação atual e risco do próximo verão. Campinas; [s.d.] [acesso em 12 abr 2012]. Disponível em: <http://2009.campinas.sp.gov.br/saude/vigilancia/informes/InformeDengueOutubro2010.pdf>.
5. Hino P, Santos CC, Santos MO, Cunha TN, Santos CB. Evolução temporal da dengue no município de Ribeirão Preto, São Paulo, 1994 a 2003. *Ciênc. saúde coletiva* 2010, 15(1): 233-238.
6. Câmara FP, Gomes AF, Santos GTD, Câmara DCP: Clima e epidemias de dengue no Estado do Rio de Janeiro. *Rev Soc Bras Med Trop* 2009, 42:137-140.
7. Lambrechts L, Paaijmans KP, Fansiri T, Carrington LB, Kramer LD, Thomas MB, Scott TW: Impact of daily temperature fluctuations on dengue virus transmission by *Aedes aegypti*. *Proc Natl Acad Sci USA* 2011, 108:7460-7465.
8. Alirol E, Getaz L, Stoll B, Chappuis F, Loutan L: Urbanisation and infectious diseases in a globalised world. *Lancet Infect Dis* 2011, 11:131-141.
9. Gubler DJ: Emerging vector-borne *Flavivirus* diseases: are vaccines the solution? *Expert Rev Vaccines* 2011, 10:563-565.
10. Eisen L, Lozano-Fuentes S: Use of Mapping and Spatial and Space-Time Modeling Approaches in Operational Control of *Aedes aegypti* and Dengue. *PLoS Negl Trop Dis* 2009, 3:7.
11. Lawson A, Biggeri A, Böhning D, Lesaffre E, Viel JF, Bertollini R: Disease mapping and risk assessment for public health. Wiley; 1999:502 p.
12. Elliott P, Wakefield J, Best N, Briggs D: *Spatial Epidemiology: Methods and Applications*. London: Oxford University Press; 2001:475p.
13. Siqueira JB, Martelli CMT, Maciel IJ, Oliveira RM, Ribeiro MG, Amorim FP, Moreira BC, Cardoso DDP, Souza WV, Andrade ALSS: Household survey of dengue infection in central Brazil: spatial point pattern analysis and risk factors assessment. *Am J Trop Med Hyg* 2004, 71:646-651.
14. Mondini A, Chiaravalloti-Neto F: Spatial correlation of incidence of dengue with socioeconomic, demographic and environmental variables in a Brazilian city. *Sci Total Environ* 2008, 393:241-248.
15. Malavasi HG. Análise espacial da epidemia de dengue em Campinas/ SP no ano de 2007.[Dissertação de Mestrado]. Campinas: Faculdade de Ciências Médicas, Universidade Estadual de Campinas; 2011.
16. Mondini A, Chiaravalloti Neto F: Socioeconomic variables and dengue transmission. *Cad Saúde Pública* 2007, 41:923-930.

17. Resendes APC, Silveira NAPR, Sabroza PC, Souza-Santos R: Determinação de áreas prioritárias para ações de controle da dengue. *Rev Saúde Pública* 2010, 44:274-282.
18. Hair Jr, Black WC, Babin, BJ, Anderson R.E, Tatham RL. *Multivariate Data Analysis*. 6.ed. Upper Saddle River, NJ: Pearson Prentice Hall, 2006: 730p.
19. Besag J, York J, Molli A: Bayesian image restoration, with two applications in spatial statistics. *Ann Inst Statist Math* 1991, 43:1-20.
20. Fahrmeir L, Lang S: Bayesian inference for generalized additive mixed models based on Markov random field priors. *J R Stat Soc Ser C Appl Stat* 2001, 50:201-220.
21. Kelsall J, Wakefield J: Discussion of “Bayesian models for spatially correlated disease and exposure data” In: Bernardo J, Berger J, Dawid A, Smith A (Ed) *Bayesian Statistics 6*. Oxford, UK: Oxford University Press; 1999:131-156.
22. Knorr-Held L, Rue H: On Block Updating in Markov Random Field Models for Disease Mapping. *Scandinavian Journal of Statistics* 2002, 29:597-614.
23. Spiegelhalter DJ, Best NG, Carlin BP, Van Der Linde A: Bayesian measures of model complexity and fit. *J R Stat Soc Series B Stat Methodol* 2002, 64:583-639.
24. Rue H, Martino S, Chopin N: Approximate Bayesian inference for latent Gaussian models by using integrated nested Laplace approximations. *J R Stat Soc Series B Stat Methodol* 2009, 71:319-392.
25. Richardson S, Thomson A, Best N, Elliott P: Interpreting Posterior Relative Risk Estimates in Disease-Mapping Studies. *Environ Health Perspect* 2004, 112:1016-1025.
26. Cordeiro R, Donalisio MR, Andrade VR, Mafra AC, Nucci LB, Brown JC, Stephan C: Spatial distribution of the risk of dengue fever in southeast Brazil, 2006-2007. *BMC Public Health* 2011, 11:355.
27. Kuno G: Factors influencing the transmission of dengue viruses. In *Dengue and Dengue Hemorrhagic Fever*. edited by Gubler DJ, Kuno G CAB International; 1997:61-88.
28. Guzmán MG, Kourí G: Dengue: an update. *Lancet Infect Dis* 2002, 2:33-42.
29. Almeida ASD, Medronho RDA, Valencia LIO: Spatial analysis of dengue and the socioeconomic context of the city of Rio de Janeiro (Southeastern Brazil). *Rev Saúde Pública* 2009, 43:666-673.
30. Pessanha JEM, Caiaffa WT, KrooLn EG, Proietti FA: Dengue em três distritos sanitários de Belo Horizonte, Brasil: inquérito soropidemiológico de base populacional, 2006 a 2007. *Rev Panam Salud Publica* 2010, 27:252-258.
31. Nagao Y, Svasti P, Tawatsin A, Thavara U: Geographical structure of dengue transmission and its determinants in Thailand. *Epidemiol Infect* 2008, 136:843-51.

32. Thammapalo S, Chongsuvivatwong V, Geater A, Dueravee M: Environmental factors and incidence of dengue fever and dengue haemorrhagic fever in an urban area, Southern Thailand. *Epidemiol Infect* 2008, 136:135-143.
33. Spiegel JM, Bonet M, Ibarra A-M, Pagliccia N, Ouellette V, Yassi A: Social and environmental determinants of *Aedes aegypti* infestation in Central Havana: results of a case-control study nested in an integrated dengue surveillance programme in Cuba. *Trop Med Int Health* 2007, 12:503-510.
34. Reiter P, Lathrop S, Bunning M, Biggerstaff B, Singer D, Tiwari T, Baber L, Amador M, Thirion J, Hayes J, Seca C, Mendez J, Ramirez B, Robinson J, Rawlings J, Vorndam V, Waterman S, Gubler D, Clark G, Hayes E: Texas Lifestyle Limits Transmission of Dengue Virus. *Emerg Infect Dis* 2003, 9:86-89.
35. Teixeira TRA, Cruz OG: Spatial modeling of dengue and socio-environmental indicators in the city of Rio de Janeiro, Brazil. *Cad Saúde Pública* 2011, 27:591-602.
36. Machado JP, Oliveira RMD, Souza-Santos R: Spatial analysis of dengue occurrence and living conditions in Nova Iguaçu, Rio de Janeiro State, Brazil. *Cad Saúde Pública* 2009, 25:1025-1034.
37. Hu W, Clements A, Williams G, Tong S, Mengersen K: Spatial patterns and socioecological drivers of dengue Fever transmission in Queensland, Australia. *Environ Health Perspect* 2012, 120:260-266.
38. Mondini A, Bronzoni RVM, Nunes SHP, Chiaravalloti Neto F, Massad E, et al. (2009) Spatio-Temporal Tracking and Phylodynamics of an Urban Dengue 3 Outbreak in São Paulo, Brazil. *PLoS Negl Trop Dis* 3(5):e448. doi:10.1371/journal.pntd.0000448
39. Flauzino RF, Souza-Santos R, Barcellos C, Gracie R, Magalhães MDAFM, Oliveira RMD: Spatial heterogeneity of dengue fever in local studies, City of Niterói, Southeastern Brazil. *Cad Saúde Pública* 2009, 43:1035-1043.
40. Organização Mundial de Saúde (OMS): Dengue guidelines, for diagnosis, treatment, prevention and control. New edition. Geneva: 2009. (disponível em : whqlib.doc.who.int/publications/2009/9789241547871_eng.pdf, acessado em 06/Abr/2012).
41. Teixeira TRA, Medronho RA: Indicadores sócio-demográficos e a epidemia de dengue em 2002 no Estado do Rio de Janeiro, Brasil. *Cad Saúde Pública* 2008, 24:2160-2170.
42. Openshaw S: Ecological fallacies and the analysis of areal census data. *Environ Plann A* 1984, 16:17-31.

4.2. Artigo 2

MODELAGEM ESPACIAL DO RISCO DE DENGUE: UMA APLICAÇÃO DE MODELOS HIERÁRQUICOS BAYESIANOS PARA DADOS COM SUPERDISPERSÃO

José Vilton Costa¹,
Liciane Vaz de Arruda Silveira²,
Maria Rita Donalisio Cordeiro¹

¹Faculdade de Ciências Médicas - UNICAMP. Departamento de Saúde Coletiva

²Instituto de Biociências – UNESP/Botucatu

RESUMO

A análise de dados de contagem para pequenas áreas geralmente viola as suposições dos modelos tradicionais de Poisson, devido à superdispersão e excesso de zeros. O excesso de zeros nos dados representa um grande desafio à modelagem de dados para pequenas áreas. A agregação de áreas com o intuito de diminuir a dispersão dos dados pode resultar em estimativas tendenciosas de risco, sendo necessário buscar soluções adequadas para análise dos dados no nível mais desagregado. O objetivo deste estudo foi analisar a relação entre as condições socioambientais e o risco de incidência de dengue segundo a abordagem bayesiana. O excesso de zero e superdispersão foram ajustados por modelos de mistura, como binomial negativo, e modelos inflacionado de zero. Os resultados não demonstraram associação entre o risco de dengue e condições socioambientais e o modelo de Poisson inflacionado de zeros com efeitos espacialmente estruturados e não estruturados apresentou melhor desempenho em relação aos demais modelos.

Palavras chaves: Análise Espacial; Dengue; Estudos Ecológicos; Modelos Bayesianos de Decisão Estatística

ABSTRACT

The analysis of count data for small areas usually violates the assumptions of the traditional Poisson models, due to overdispersion and excess zeros. The excess zeros in the data represents a major challenge to modeling data for small areas. The aggregation of areas in order to reduce the dispersion of data can result in biased estimates of risk, the necessity to seek appropriate solutions for data analysis in more disaggregated level. The aim of this study was to analyze the relationship between socio-environmental conditions and risk of incidence of dengue according to the Bayesian approach. An excess of zeros and overdispersion was adjusted with by estimating mixture models such as negative binomial and, zero-inflated models. The results showed no association between the risk of dengue and socio-environmental conditions and zero-inflated Poisson model with spatially structured and unstructured effect performed better than the other models.

Keywords: Spacial Analysis; Dengue; Ecological Estudios; Bayesian Statistical Models of Decision

Introdução

Dengue é uma doença viral de transmissão vetorial causada por um dos quatro sorotipos do vírus dengue (DENV-1, DENV-2, DENV3 e DENV-4), pertencentes ao gênero *Flavivirus*, da família *Flaviviridae*. A Organização Mundial da Saúde (OMS) considera o dengue um dos maiores problemas de saúde pública, com cerca de 100 milhões de casos anualmente no mundo, sendo que mais de 2,5 bilhões de pessoas vivem em área sob risco de infecção¹.

Atualmente não se dispõe de uma vacina eficaz para uso preventivo contra dengue. Na ausência de uma vacina eficaz, as estratégias de prevenção e controle são direcionadas ao combate direto ao mosquito vetor, considerado o único elemento controlável da cadeia epidemiológica da dengue². Assim, uma análise dos casos de dengue registrados e uma identificação dos padrões de risco se tornam necessárias. Com esta informação é possível desenvolver estratégias de prevenção e um controle eficaz de epidemias.

A incidência de inúmeras doenças infecciosas, particularmente aquelas transmitidas por vetores, geralmente exibem padrões espaciais, as quais podem ser resultantes das interações entre patógeno-hospedeiro-ambiente, em locais específicos e em períodos específicos⁴⁻⁶. A natureza, intensidade, direção e o resultado dessas interações específicas dependem da forma como as populações envolvidas ocupam e se apropriam do espaço^{7,8}.

A disponibilidade de dados geograficamente indexados de saúde e população, os avanços da computação, dos sistemas de informação geográfica (SIG) e de metodologias estatísticas, abriram caminho para a exploração de estatísticas de saúde para pequenas áreas com base em dados de rotina. Essas análises podem ser usadas para tratar de questões específicas relativas à saúde em relação a fatores ambientais, em investigações de agrupamento de doenças ou para a geração de hipóteses^{3,9}.

A estimativa das taxas de incidência e a identificação de fatores associados a sua distribuição são fundamentais para que medidas adequadas de vigilância possam ser tomadas visando à prevenção e o controle dessas doenças. O estudo da variação geográfica das taxas de doenças em pequenas áreas apresenta algumas características que devem ser

consideradas na estimação dos modelos empíricos¹⁰. Em geral, assume-se que a contagem de casos ou a taxa de ocorrência de uma doença segue uma distribuição de Poisson³.

O modelo de Poisson assume que a média e a variância são iguais. No entanto, quando a incidência é baixa, a distribuição espacial da doença exibe uma grande quantidade de áreas com contagens zeros. Quando isso ocorre, normalmente também surgem violações as suposições básicas das distribuições de probabilidade. Um destes casos é a superdispersão^{11,12}, que ocorre quando a variância é consideravelmente maior que a média. Não considerar esta superdispersão pode levar à subestimativas de erros-padrão e a uma inferência deficiente dos parâmetros da regressão¹².

A distribuição Binomial Negativa (BN) é comumente utilizada para modelar dados com superdispersão em substituição à distribuição de Poisson. O modelo BN é uma mistura de uma distribuição de Poisson e uma distribuição Gama, que considera a superdispersão como resultado de uma heterogeneidade não observada entre as observações. A superdispersão é ajustada à medida que a variância condicional passa a apresentar uma forma funcional quadrática ou linear da média condicional¹².

A superdispersão nos dados também pode ser decorrente do excesso de zeros (isto é, a frequência observada de zeros excede em muito ao número esperado, segundo as premissas da distribuição dos dados). Apesar de minimizar a superdispersão, o modelo BN é insuficiente para controlar a presença de excesso de zeros. Para superar esta limitação, foram desenvolvidos modelos inflacionados de zeros. Revisões prévias destes métodos incluem os trabalhos de Ridout et al.¹³, Hall¹⁴ e Gschlöbl e Czado¹⁵ entre outros. Esses modelos, cujas aplicações se destacam a partir do trabalho de Mullahy¹⁶, têm sido usados com sucesso em econometria, demografia, medicina, saúde pública, epidemiologia, biologia e em muitos outros campos. Os modelos inflacionados de zeros caracterizam-se por ajustar os dados a partir de uma mistura de duas populações: uma caracterizada pela contagem de zeros e outra em que as contagens são realizações de uma distribuição discreta (por exemplo, Poisson e BN). Desta forma, podem ser empregado em conjunto com qualquer uma das distribuições para dados de contagem¹⁷.

Neste artigo, consideramos o modelo de Poisson Inflacionado de Zeros (ZIP)¹⁸ e o modelo Binomial Negativo inflacionado de zeros (ZINB)¹⁹. Os modelos ZIP e ZINB, uma

mistura entre as distribuições de Poisson/ Binomial negativa e uma distribuição de Bernoulli, permitem controlar a superdispersão decorrente do excesso de valores zeros nos dados, com o objetivo de inflacionar a probabilidade da ocorrência de valores zero, através da inclusão de uma variável aleatória latente que representa a ausência ou presença do processo em observação^{18,19}.

Adicionalmente, na análise de dados agregados espacialmente, a superdispersão também pode ocorrer em virtude de efeitos espaciais, a saber, autocorrelação espacial e heterogeneidade espacial²⁰. Adotando-se a abordagem Bayesiana hierárquica²¹, a estrutura de dependência espacial é modelada a partir da inclusão de efeitos aleatórios no modelo, sob uma estrutura autoregressivo condicional (CAR)²². Os modelos bayesianos destacam-se por apresentar boa capacidade para ajustar a variabilidade das estimativas para pequenas áreas geográficas, elucidando tendências globais, padrões geográficos, além de possibilitar a inclusão de covariáveis. Agawal *et al.*²³ propuseram um modelo ZIP para modelar a distribuição espacial de uma espécie de crustáceo terrestre. Musal e Aktekin²⁴, utilizando a abordagem bayesiana espacial, aplicaram o modelo ZIP em combinação com efeitos aleatórios para investigar o efeito das condições socioeconômicas sobre a mortalidade por HIV. Gschlöbl e Czado¹⁵, por sua vez, compararam diferentes modelos de contagem em conjunção com a inclusão de efeitos aleatórios para acomodar a superdispersão presente no conjunto de dados analisados sobre a distribuição espacial da doença meningocócica invasiva.

O objetivo deste artigo foi analisar a associação entre condições socioambientais e risco de dengue na cidade de Campinas, a partir da aplicação de modelos de regressão espacial bayesianos para dados de contagem, ajustando a superdispersão e o excesso de zeros presente nos dados.

Métodos

O estudo foi realizado na cidade de Campinas, Estado de São Paulo. A cidade está situada na região sudoeste do Estado (47°04'40'' Longitude Oeste e 22°53'20'' Latitude Sul), numa altitude média de 640 metros acima do nível do mar. A população

estimada em 2011, segundo o Instituto Brasileiro de Geografia e Estatística (IBGE) era de 1.088.611 habitantes.

Foram utilizados os dados referentes aos casos de dengue autóctones, registrados entre 01 de janeiro e 31 de dezembro de 2007, ano de maior incidência de dengue na cidade. O banco de dados com os casos foram obtidos do Sistema de Informação de Agravos de Notificação (SINAN), fornecido pela Coordenadoria de Vigilância e Saúde da Secretária Municipal de Saúde do município (Covisa/SMS-Campinas). As unidades geográficas de análise do estudo foram os setores censitários definidos pelo Censo Demográfico de 2000, do IBGE. A malha digital dos setores e os dados com a população de cada setor, referentes ao Censo Demográfico 2000, foram obtidos no site do IBGE (IBGE; <http://www.ibge.gov.br>).

Foi criada uma variável sintética para caracterização do nível de condições socioambientais (ICSA) dos setores censitários, segundo três classes de carência socioambiental: maior; moderada e menor. Esta variável foi obtida a partir dos seguintes indicadores do Censo 2000: proporção de pessoas responsáveis por domicílios com renda até três salários-mínimos; proporção de pessoas responsáveis por domicílios com até três anos de estudo; proporção de domicílios com esgotamento sanitário inadequado e proporção de domicílios com destinação do lixo inadequada. Estes variáveis foram submetidas à técnica de componentes principais (ACP). O primeiro componente principal da ACP foi analisado pelo método estatístico de análise de agrupamentos (*cluster analysis*), resultando no indicador sintético com três categorias.

Modelos bayesianos hierárquicos foram utilizados para estimar o risco relativo de dengue nos setores censitários. Considerando-se que a área de estudo (município de Campinas) é dividida em N áreas adjacentes (1313 setores censitários) e θ_i representa o risco relativo para o i -ésimo setor censitário ($i = 1, \dots, n$). Além disto, Y_i e E_i representam a contagem dos casos de dengue (variável resposta) observada e os casos esperados no i -ésimo setor, respectivamente. O modelo clássico assume que Y_i segue uma distribuição de Poisson, com $Y_i | \theta_i \sim \text{Poisson}(\mu_i = E_i \theta_i)$ ^{20,25}. A estimativa de interesse é o risco relativo θ_i de cada sub-área i , em que o estimador de máxima verossimilhança para θ_i coincide com

a razão de incidência padronizada (SIR - *standardized incidence rate*), definida por:
 $\hat{\theta}_i = SIR_i = Y_i/E_i$.

Os modelos de contagem BN, ZIP e ZINB foram utilizados para ajustar a superdispersão e o excesso de zeros. Esses modelos apresentam as seguintes especificações:

$$Y_i | \mu_i, \delta \sim BN(\mu_i, \delta) \quad (1)$$

$$Y_i | \mu_i, \omega_i \sim ZIP(\mu_i, \omega_i) \quad (2)$$

$$Y_i | \mu_i, \delta, \omega_i \sim ZINB(\mu_i, \delta, \omega_i) \quad (3)$$

Nos modelos, descritos em (1-3), μ_i é a média de uma Poisson, δ é o parâmetro de superdispersão e ω_i é a proporção de zeros do modelo.

Os modelos inflacionados de zeros (ZIP e ZINB) surgem quando a variável resposta é modelada como uma mistura de uma distribuição de Bernoulli para a presença/ausência de zeros e outra distribuição para os valores positivos (Poisson ou BN). O excesso de zeros é ajustado assumindo que os zeros surgem a partir de dois distintos processos. O primeiro processo ocorre com probabilidade ω_i e produz apenas verdadeiros zeros (ou zeros estruturais), o segundo, denominado falsos zeros (ou zeros amostrais) ocorre com probabilidade $(1 - \omega_i) \Pr(Y_i = 0)$, em que $\Pr(Y_i = 0)$ é uma distribuição de Poisson ou BN. A probabilidade total de zeros é a combinação das probabilidades dos dois processos, ou seja, $\Pr(Y_i = 0) = \omega_i + (1 - \omega_i) \Pr(Y_i = 0)^{26}$.

Neste estudo, os zeros verdadeiros representam a situação em que não houve registros de casos em um dado setor, mesmo com a presença de uma população humana suscetível e com a presença do vetor e/ou patógeno no ambiente. Isto pode ser devido à resistência imunológica da população ou porque a transmissão não ocorreu. Ou seja, a situação era favorável para a incidência da doença, mas ela não ocorreu. Os falsos zeros representam a situação em que a população não se encontrava suscetível a doença em virtude da ausência do vetor e/ou patógeno no ambiente ou a situação em que realmente houve ocorrência de casos de dengue em um dado setor, mas nenhum caso foi registrado, possivelmente por subnotificação²⁷.

As contagens positivas (isto é, maiores do que zero) ocorrem com probabilidade $(1 - \omega_i) \Pr(Y_i = y_i)$, em que $\Pr(Y_i = y_i)$ é uma função de probabilidade de Poisson ou BN para as contagens positivas.

Segundo o modelo bayesiano proposto por Besag, York e Mollié²⁸, denotado por modelo BYM, foram incluídos efeitos aleatórios nos modelos de regressão para contornar a heterogeneidade espacial, bem como, autocorrelação espacial dos dados. Considerando-se para estes modelos a estrutura dos modelos lineares generalizados (GLM)²⁹, assumindo-se a função de ligação logito entre o risco de dengue θ_i e o efeito da covariável ICSA, os preditores lineares do risco relativo θ_i e o modelo de excesso de zeros são definidos por:

$$\log(\theta_i) = \alpha + \beta_1 \text{Moderada}_i + \beta_2 \text{Maior}_i + fu_i + fs_i \quad (4)$$

$$\log\left(\frac{\omega_i}{1 - \omega_i}\right) = \alpha + \beta_1 \text{Moderada}_i + \beta_2 \text{Maior}_i + fu_i + fs_i \quad (5)$$

$$i=1, \dots, n$$

em que n é o total de setores censitários, a equação (4) representa o modelo de médias do risco relativo para os modelos de regressão de Poisson, BN, ZIP e ZINB. A equação (5) modela o excesso de zeros nos modelos ZIP e ZINB, segundo um modelo logístico. Nestas duas equações (4) e (5), o parâmetro α é o intercepto e representa o logaritmo do risco médio global; β_1 e β_2 são os efeitos das categorias da covariável ICSA, assumindo-se a categoria “menor” como referência; fu_i e fs_i são vetores de efeitos aleatórios que modelam a heterogeneidade não estruturada (independente) e espacialmente estruturada, respectivamente²⁵.

Segundo o paradigma bayesiano, a dependência espacial é introduzida assumindo-se uma *priori* para fs_i com estrutura autorregressiva condicional (CAR), definida por:

$$\gamma_i | \gamma_j, \tau_s \sim N\left(\frac{1}{n_i} \sum_{j \in \partial_i} \gamma_j, \frac{1}{n_i \tau_s}\right) \quad (6)$$

em que n_i corresponde ao número de setores vizinhos (adjacentes) e $j \in \partial_i$ denota que o setor j é vizinho do setor i . Assim, a média (condicional) de γ_i é uma média não ponderada. A especificação da distribuição CAR depende de um único parâmetro $\tau_s = 1/\sigma_s^2$, em que τ_s e σ_s^2 representam a precisão e a variabilidade condicional espacialmente estruturada, respectivamente, e $N(a,b)$ é anotação genérica para uma distribuição normal com média a e variância b .

O efeito não estruturado fu_i foi modelado por uma distribuição normal com média zero e parâmetro de precisão τ_u . A distribuição adotada para τ_s e τ_u foi uma densidade Gama pouco informativa, denotada por $\tau_s, \tau_u \sim \text{Gama}(0,5; 0,0005)$, conforme recomendação da literatura²⁰. Para os parâmetros α e β_p foram atribuídas *prioris* com distribuição normal com média igual a zero e parâmetro de precisão o menor possível.

Neste artigo foram ajustados doze diferentes modelos, correspondendo a três diferentes formulações para cada uma das distribuições de Poisson, BN, ZIP e ZINB. Em cada uma dessas distribuições, os modelos ajustados distinguem-se em relação à presença de efeito não estruturado ($fu = fu_i$), efeito espacial estruturado ($fs = fs_i$) e efeito fixo da covariável de condições socioambientais (ICSA). Segundo a presença/ ausência destes efeitos, a formulação básica dos modelos ajustados em cada uma das distribuições foi:

$$a) \log(\theta_i) = \alpha + fu$$

$$b) \log(\theta_i) = \alpha + fu + fs$$

$$c) \log(\theta_i) = \alpha + fu + fs + \text{ICSA}$$

O método INLA (*Integrated Nested Laplace Approximation*)³⁰ foi utilizado para estimativa dos parâmetros. A abordagem INLA foi desenvolvida para permitir inferência Bayesiana acurada, a partir da classe de modelos que envolvem Campos Aleatórios Markovianos Gaussianos (CAMG). Destacando-se por empregar aproximações de Laplace para aproximar as *posteriors* via métodos computacionais determinísticos, sem a necessidade de utilizar simulações. Assim, as estimativas dos parâmetros são computacionalmente mais rápidas e eficientes em relação aos métodos tradicionais de

Monte Carlo via cadeias de Markov (MCMC), sem o inconveniente de realizar diagnóstico de convergência de cadeias simuladas³¹. As análises foram realizadas com o software R, versão 2.15.1 (R Development Core Team, 2012), utilizando o pacote R-INLA³².

Os modelos foram comparados usando o critério de informação da *Deviance* (DIC)³¹ e a *Deviance* baseada na Ordenada Preditiva Condicional (CPO - *Conditional Predictive Ordinate*) para cada observação^{34,35}. As CPO's são densidades de validação cruzada, isto é, distribuições preditivas condicionadas sobre o conjunto de dados observados exceto a *k*-ésima informação, proporcionando uma medida de ajuste para cada observação individualmente. Comparações entre os modelos são realizadas através da média do logaritmo do CPO (-média(log(CPO)) ou \overline{LCPO}). Em ambos os casos, o menor DIC ou menor \overline{LCPO} , representam o modelo com ajuste mais adequado. Adicionalmente, estimou-se e mapeou a probabilidade *a posteriori* do risco relativo (RR) de dengue ser maior do que um ($\Pr[RR > 1 | Y]$), obtendo-se assim, uma medida de evidência (estatística) de excesso de risco em cada área. Estimativas de probabilidades *a posteriori* superiores a 0,8 representam evidências de excesso de risco nas áreas³⁶. A distribuição espacial destas áreas de risco permite avaliar possível efeito de agregação espacial destas áreas.

Resultados

Um total de 11.519 casos de dengue foi confirmado e notificado³⁵ ao SINAN, na cidade de Campinas, em 2007. Destes, 8982 foram georreferenciados sobre a malha digital dos 1313 setores censitários do município. Entretanto, no processo de modelagem, treze setores foram excluídos em virtude da ausência do total da população residente nos dados populacionais do Censo 2000, ocasionando a exclusão de 60 casos nestes setores. O banco de dados final para modelagem foi constituído de 8922 casos de dengue.

O mapa da Figura 1 apresenta a distribuição espacial da covariável ICSA, em que se observa uma grande heterogeneidade na distribuição espacial das condições socioambientais em relação aos setores censitários. Em termos quantitativos, 47,0% dos setores apresentaram menor carência, 42,5% carência moderada e 10,5% maior carência socioambiental.

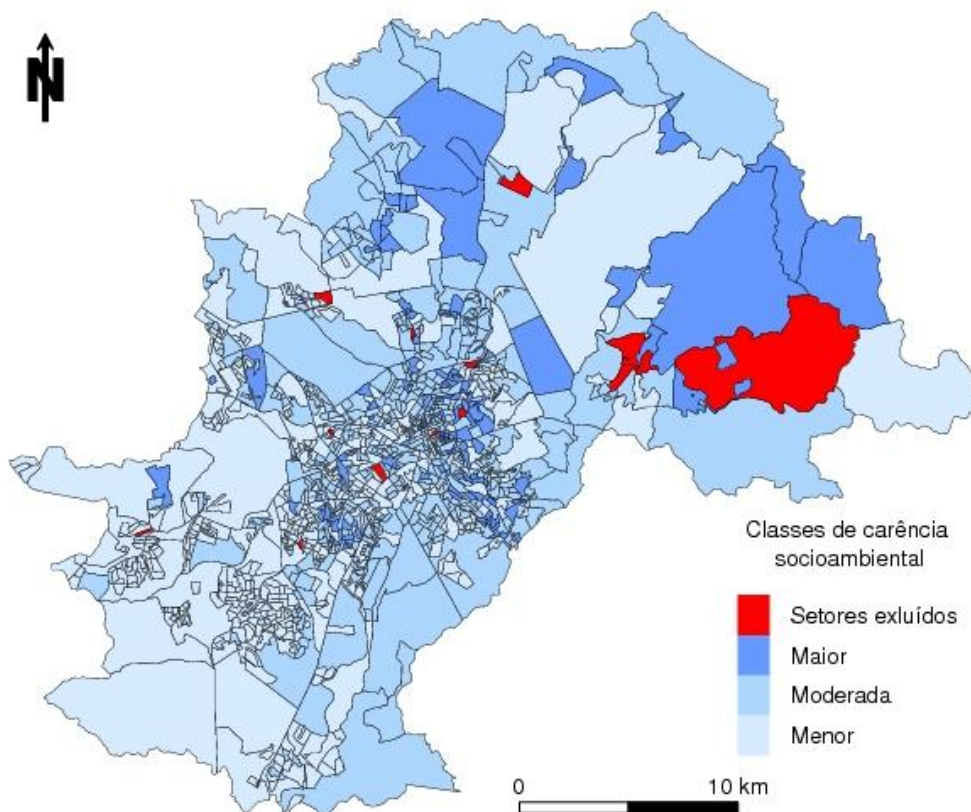


Figura 1. Distribuição espacial do índice de carência socioambiental (ICSA), segundo o setor censitário.

Na Figura 2, observa-se a distribuição de frequência das contagens dos casos de dengue. O histograma mostra alta variação na distribuição de frequência, caracterizada por uma assimetria positiva e com excesso de áreas sem registros de casos da doença. Verificou-se que em 21% dos setores censitários não foi registrado nenhum caso de dengue. A média e o desvio padrão do número de casos de dengue e da taxa de incidência padronizada (SIR), por setor censitário, foram de 6,8 (12,2) casos e 1,29 (3,5), respectivamente. Estas estatísticas descritivas, em que a variância das contagens dos casos é maior que a sua média, juntamente com a distribuição de frequência presente no histograma da Figura 1, sugerem a presença de superdispersão nos dados de contagens de dengue.

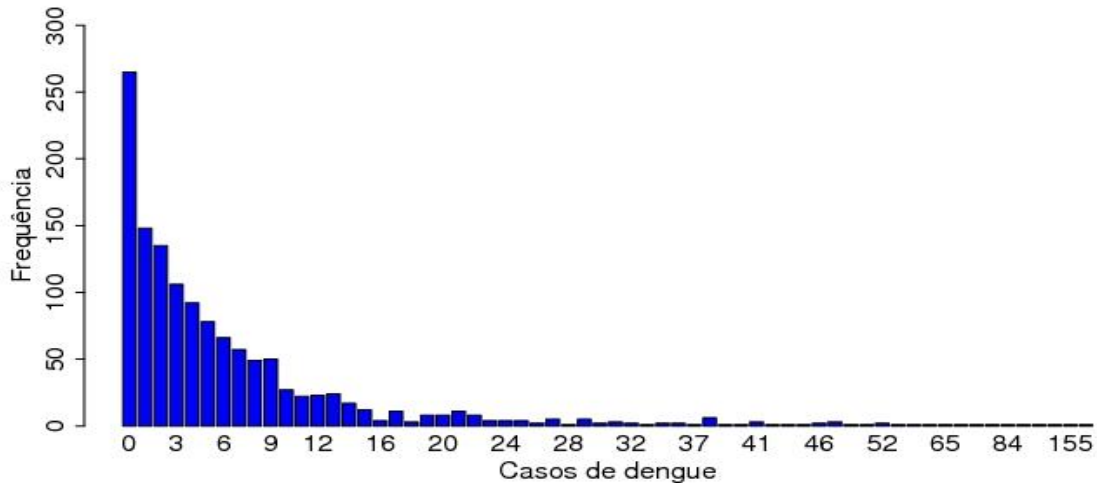


Figura 2. Distribuição de frequência dos casos de dengue.

A Tabela 1 apresenta os valores do DIC e do logaritmo do escore para o ajuste dos diferentes modelos. Considerando-se os modelos de regressão de Poisson, BN, ZIP e ZINB, o modelo ajustado somente com o efeito não estruturado fu , foi o modelo que apresentou o maior valor do DIC em cada um destes modelos de regressão. A inclusão do efeito espacial fs ajustou-se adequadamente nos quatro modelos de regressão, em que se observou diminuição nos valores do DIC em relação ao modelo com o efeito fu . A inclusão da covariável ICSA nos modelos mostrou-se adequada apenas no modelo BN. Entre todos os doze modelos ajustados, o menor valor do DIC (6551,95) foi observado no modelo ZIP com efeitos fu e fs , sendo considerado o modelo com melhor ajuste dos dados. O escore do logaritmo forneceu resultados para o ajuste do modelo semelhantes aos obtidos com o DIC.

Tabela 1. Comparação dos modelos segundo o Critério de Informação da *Deviance* (DIC) e média do Logaritmo da Ordenada Preditiva Condicional (\overline{LCPO}).

Modelos	DIC				\overline{LCPO}			
	Poisson	BN	ZIP	ZINB	Poisson	BN	ZIP	ZINB
fu	7785,48	7276,13	6719,61	6850,58	5,38	2,97	2,37	2,40
$fu+fs$	7689,81	7201,28	6551,95	6740,8	5,26	2,84	2,33	2,37
$fu+fs+ICSA$	7689,92	7195,71	6552,12	6744,42	5,25	3,04	2,33	2,36

Tabela 2. Resumo das estimativas dos parâmetros de superdispersão ($\hat{\delta}$) e da probabilidade de valores zero ($\hat{\omega}$).

Modelos	Superdispersão ($\hat{\delta}$)		Probabilidade de valores zero ($\hat{\omega}$)	
	BN	ZINB	ZIP	ZINB
fu	0,544 (0,500; 0,591)	0,554 (0,507; 0,604)	0,181 (0,163; 0,208)	0,008 (0,001; 0,035)
fu+fs	0,591 (0,537; 0,648)	0,599 (0,542; 0,661)	0,186 (0,166; 0,201)	0,010 (0,001; 0,040)
fu+fs+ICSA	0,593 (0,542; 0,648)	0,604 (0,546; 0,667)	0,183 (0,164; 0,202)	0,010 (0,001; 0,042)

Na Tabela 2, podemos avaliar complementarmente a capacidade dos modelos ajustarem a superdispersão presente nos dados. O parâmetro δ , descritos em (1) e (3) captura a variação extra-Poisson presente nos dados através das distribuições BN e ZINB^{12,19}, mostrando que quanto maior o valor de δ menor a variabilidade dos dados³⁵. Observou-se em cada um destes modelos de distribuição que a estimativa do parâmetro δ aumentou com o acréscimo do efeito fs , diminuindo conseqüentemente a variabilidade dos dados. A inclusão da covariável ICSA aos modelos especificados previamente com os efeitos fu e fs produziu pouca alteração no parâmetro δ de superdispersão. Os valores estimados de δ em cada um dos modelos ZINB foram semelhantes aos ajustados com o modelo BN. Conforme descrito em (2) e (3), a probabilidade de excesso de zeros ajustada pelos modelos ZIP e ZINB e avaliada pelo parâmetro ω , mostrou que a inclusão do efeito fs e da covariável ICSA apresentaram pouca contribuição para capturar o excesso de zeros. Segundo as estimativas do parâmetro ω , a proporção estimada de valores zero nos dados foi maior para os modelos ZIP (18,6%) em relação ao modelo ZINB (1%).

A Tabela 3 resume os valores do DIC e do logaritmo do escore para o ajuste dos diferentes modelos. Considerando-se os modelos de regressão de Poisson, BN, ZIP e ZINB, o modelo ajustado somente com o efeito não estruturado fu , foi o modelo que apresentou o maior valor do DIC em cada um destes modelos de regressão. A inclusão do efeito espacial fs mostrou-se adequado nos quatro modelos de regressão, em que se observa diminuição nos valores do DIC em relação ao modelo com o efeito fu . Por outro lado, o efeito da covariável ICSA mostrou-se adequado apenas no modelo BN. Entre todos os doze modelos ajustados, o menor valor do DIC (6551,95) foi observado no modelo ZIP com efeitos fu e fs , sendo considerado o modelo com melhor ajuste dos dados. O escore do logaritmo forneceu resultados para o ajuste do modelo semelhantes aos obtidos com o DIC.

Tabela 3. Comparação dos modelos segundo o Critério de Informação da *Deviance* (DIC) e média do Logaritmo da Ordenada Preditiva Condicional (\overline{LCPO}).

Modelos	DIC				\overline{LCPO}			
	Poisson	BN	ZIP	ZINB	Poisson	BN	ZIP	ZINB
fu	7785,48	7276,13	6719,61	6850,58	5,38	2,97	2,37	2,40
fu+fs	7689,81	7201,28	6551,95	6740,8	5,26	2,84	2,33	2,37
fu+fs+ICSA	7689,92	7195,71	6552,12	6744,42	5,25	3,04	2,33	2,36

Na Figura 3, mapeou-se (a) a razão de incidência padronizada (SIR), (b) as estimativas *a posteriori* da média do risco relativo do modelo que melhor ajustou-se aos dados, ou seja, o modelo ZIP com os efeitos aleatórios *fu* e *fs* e (c) a probabilidade *a posteriori* do risco relativo ser maior do que um. As estimativas *a posteriori* do risco relativo variaram de 0,20 a 55,3, com valor médio de 1,41 e desvio padrão de 2,3. Contrastando-se as distribuição espacial da SIR com a distribuição do risco relativo (mapas a e b), nota-se que há uma redução de áreas com risco relativo próximo à zero e um estreitamento das estimativas do risco em direção ao risco médio, em função do efeito de suavização do modelo espacial. Observou-se que em 566 (43,6%) dos setores o risco relativo foi acima de 1 (um) e segundo as estimativas de probabilidade *a posteriori* (mapa c), 264 (21%) dos setores analisados apresentaram excesso de risco, com probabilidade acima de 0,8 do risco relativo ser maior do que um ($RR > 1$). Nota-se que os setores com excesso de risco não se encontram distribuídas de forma aleatória, pelo contrário, formam agregados de área.

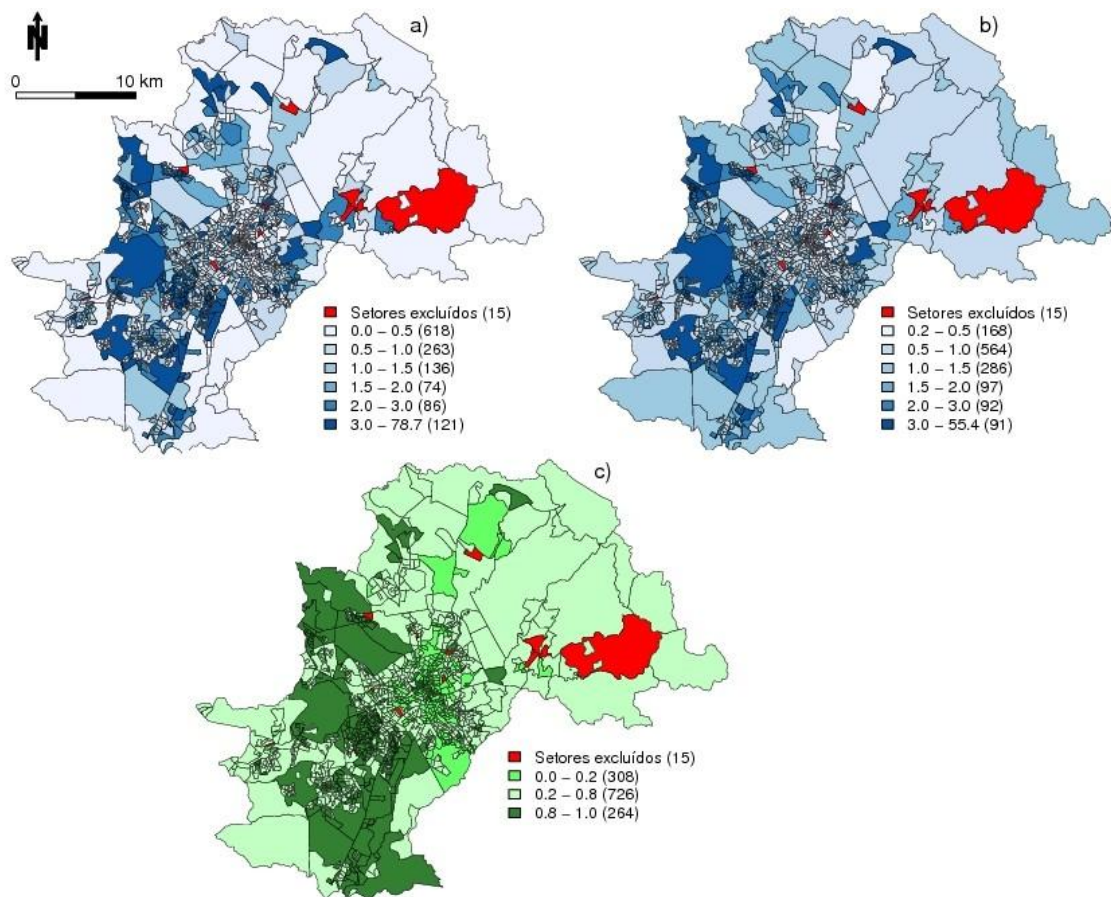


Figura 3. Mapas do risco relativo para os dados de dengue: a) razão de incidência padronizada (SIR); b) estimativa a posteriori da média do risco relativo para o modelo ZIP, com efeitos não estruturado (f_u) e espacialmente estruturado (f_s) e c) probabilidade a posteriori.

Discussão e conclusões

Este estudo utilizou a abordagem de modelagem bayesiana para o desenvolvimento da análise espacial do risco de dengue em escala local. Modelos de regressão para dados de contagem foram empregados com o objetivo de elucidar o risco de incidência de dengue em relação a condições socioambientais, identificadas em estudos prévios³⁸⁻⁴⁴.

Um problema presente na análise de dados agregados em pequenas unidades geográficas diz respeito à presença de superdispersão e excesso de zeros. A aplicação dos

modelos de contagem BN, ZIP e ZINB, alternativos ao modelo tradicional de Poisson, permitiu reduzir o efeito de instabilidade, ajustando a superdispersão e o excesso de zeros presente nos dados agregados por setor censitário. Adicionalmente, efeitos aleatórios espacialmente estruturados e não estruturados foram modelados a partir de *prioris* definidas por um campo aleatório de Markov e *prioris* Gaussianas com média zero, respectivamente²⁵. O método INLA foi adotado para estimação dos parâmetros do modelo de regressão.

Os resultados obtidos demonstraram a presença de variação na distribuição geográfica na incidência de dengue na cidade de Campinas, e esta variação não se mostrou associada às características socioambientais das áreas analisadas. Sendo essa variação espacial explicada por outros fatores não mensurados. Detectou-se um conjunto de áreas com excesso de risco, em que é crucial para os sistemas de vigilância em saúde a identificação destas áreas para o planejamento de suas ações de prevenção e controle do dengue.

Estudos sobre a associação entre fatores socioambientais e incidência de dengue apresentam resultados contraditórios⁴⁰. Almeida *et al.*³⁹ na cidade do Rio de Janeiro, Pessanha *et al.*⁴² em Belo Horizonte (MG) e Resendas *et al.*⁴³ na cidade de Niterói (RJ) encontraram relação positiva entre condições socioambientais e incidência de dengue. Distintamente, Mondini e Chiaravalloti-Neto³⁸ na cidade de São José do Rio Preto (SP), Machado *et al.*⁴¹ na cidade de Nova e Iguaçu (RJ) e Teixeira *et al.*⁴⁴ na cidade do Rio de Janeiro (RJ) encontraram resultados semelhantes aos encontrados no nosso estudo.

As associações divergentes encontradas nos estudos sobre a ocorrência da dengue e condições socioambientais podem estar relacionadas a diferentes estratégias metodológicas. O tipo de dado utilizado (dados primários ou secundários), diferentes variáveis adotadas para caracterização socioambiental e o tipo de agregação utilizada (setor censitário, bairro, distritos e/ou municípios) podem gerar distorções nos diferentes estudos⁴⁰.

Os setores censitários são as menores unidades de agregação disponível para os dados socioambientais, permitindo a visualização espacial mais precisa do território onde se concentram os maiores níveis de transmissão de dengue. A agregação dos dados por esta

unidade geográfica pressupõe homogeneidade máxima intra-áreas e heterogeneidade máxima inter-áreas, segundo as suas condições socioambientais.

Entre os modelos comparados (Poisson, BN, ZIP e ZINB), o modelo de Poisson inflacionado de zeros mostrou-se o mais adequado, registrando o menor valor do DIC (Tabela 2). A proporção observada de setores censitários com contagens zeros de dengue foi de 21% e o percentual ajustado segundo o modelo ZIP foi 18,6%. A inclusão dos efeitos aleatórios espaciais nos modelos ajustou a heterogeneidade espacial não observada e corrigiu a presença de correlação espacial presente entre as observações. A correlação espacial foi modelada segundo uma matriz de vizinhança binária, definida a partir relação de vizinhança entre os setores censitários. É importante considerar que diferentes especificações na matriz de vizinhança produzem diferentes estimativas na análise espacial. Os critérios de vizinhança também podem ser determinados com base no tamanho da fronteira, o tamanho das fronteiras com ou sem a presença de barreiras naturais, ou outra característica geográfica que possa interferir na ligação entre as áreas³.

Os critérios DIC e \overline{LCPO} mostraram-se adequados para escolha do modelo com melhor ajuste, apresentando resultados semelhantes. A utilização da abordagem INLA para estimativa dos parâmetros mostrou-se uma alternativa atrativa e conveniente para realização de inferência. O processo de estimação dos parâmetros não apresentou nenhum problema de convergência, estimando os parâmetros de forma bem rápida (aproximadamente 13 segundos), conforme já demonstrado em alguns estudos^{30,45}.

Vale destacar algumas limitações quando são utilizados dados secundários na análise espacial. Quanto ao tipo de agregação utilizada, os diferentes resultados obtidos segundo diferentes recortes espaciais são chamados de problema da unidade de área modificável⁴⁶. Nesse caso, mudanças na escala e nas dimensões da unidade de análise podem levar a conclusões diferentes.

Outra limitação do estudo refere-se à possibilidade de subnotificação de casos e da variabilidade da qualidade dos dados e informações registradas nas no SINAN. São registrados principalmente aqueles casos de evolução mais grave, havendo subnotificação nas formas mais leve da doença. Essas subnotificações ocorrem frequentemente por falhas do próprio sistema de informações em saúde e conhecimento insuficiente dos profissionais

de saúde sobre a vigilância e assistência a doença²⁷. Adicionalmente, os dados registrados no SINAN retratam a busca a assistência médica nos serviços públicos, representada predominantemente pela população das classes populares, não incluindo grande parte dos casos que buscam atendimento nos serviços privados, o que pode levar a distorções no conhecimento da distribuição da circulação dos vírus do dengue⁴⁷.

A perda de registros de casos em razão de inconsistências dos endereços considera-se outra limitação. A perda de dados por endereço incompleto pode ter sido seletiva, pois é possível que haja maior perda em áreas de favelas ou de ocupação recente, podendo levar à minimização da situação daquelas áreas^{48,49}.

Por fim, os estudos ecológicos por não utilizarem dados sobre indivíduos, apenas dados agregados sobre grupos de pessoas, estão sujeitos ao “viés ecológico”, que se refere à impossibilidade das associações obtidas no nível ecológico (agregado) refletirem associações no nível individual¹. Uma alternativa para reduzir esse tipo de viés seria a combinação de dados de área e uma amostra de dados individuais. Inclusive, os resultados obtidos com esta abordagem apresentariam maior poder estatístico em relação aos estudos que utilizam somente dados individuais⁵⁰. Trabalhos futuros precisam considerar o ajuste para outros fatores de exposição no nível agregado e no nível individual, além de considerar a variação espaço-temporal da incidência de dengue.

Portanto, conclui-se que é fundamental avaliar de forma adequada a distribuição discreta que melhor ajusta o conjunto de dados de contagem em análise, não se restringindo somente a distribuição de Poisson. Os modelos inflacionados de zero permitem analisar os dados de área em níveis mais desagregados, sem a necessidade de agregar áreas com baixa frequência do evento de interesse juntamente com áreas em que não houve a ocorrência do evento. A partir da abordagem bayesiana, a aplicação dos modelos inflacionados de zeros com a inclusão de efeitos aleatórios, permite obter estimativas mais robustas da distribuição espacial do risco relativo.

Referências

43. OMS. WHO. World Health Organization. Global strategy for dengue prevention and control 2010-2020 [Internet]. Global strategy for dengue prevention and control 2010-2020. [acesso em 30 nov 2010];Disponível em:
http://apps.who.int/iris/bitstream/10665/75303/1/9789241504034_eng.pdf
44. Gubler DJ: Emerging vector-borne *Flavivirus* diseases: are vaccines the solution? *Expert Rev Vaccines* 2011, 10:563-565.
3. Elliott P, Wakefield J, Best N, Briggs D. *Spatial Epidemiology: Methods and Applications*. London: Oxford University Press, 2001.
4. Hastings A. Spatial Heterogeneity and Ecological Models. *Ecology* 1990, 11;71(426-428).
5. Ostfeld RS, Glass GE, Keesing F. Spatial epidemiology: an emerging (or re-emerging) discipline. *Trends in Ecology & Evolution* 2005, 20(6):328–336.
6. Szymanski BK, Caraco T. Spatial analysis of vector-borne disease: A four species model. *Evolutionary Ecology* 1994, 8(3):299–314.
7. Silva LJ. O conceito de espaço na epidemiologia das doenças infecciosas. *Cadernos de Saúde Pública* 1997, 13(4):585–593.
8. Favier C, Schmit D, Müller-Graf CDM, Cazelles B, Degallier N, Mondet B, et al. Influence of spatial heterogeneity on an emerging infectious disease: the case of dengue epidemics. *Proc Biol Sci* 2005, 272: 1171–1177.
9. Lawson AB. *Spatial and Syndromic Surveillance for Edited by. Prevention* 2005, 3(1):97–98.
10. Greenland S, Robins J. Invited commentary: ecologic studies-biases, misconceptions, and counterexamples. *Am J Epidemiol* 1994, 139(8):747–760.
11. Wakefield J. *Disease mapping and spatial regression with count data*. Biostatistics (Oxford, England). 2007 Apr 1;8(2):158–83.
12. Cameron AC, Trivedi PK. *Regression Analysis of Count Data*. Cambridge University Press; 1998.
13. Ridout M, Demétrio CGB, Hinde J. *Models for count data with many zeros*. Proceedings from the International Biometric Conference; Cape Town. December 1998.
14. Hall DB. Zero-inflated Poisson and binomial regression with random effects: a case study. *Biometrics* 2000, 56(4):1030–1039.

15. Gschlößl S, Czado C. Modelling count data with overdispersion and spatial effects. *Statistical Papers* 2008, 49:531–552.
16. Mullahy J. Specification and testing of some modified count data models. *Journal of Econometrics* 1986, 33(3):341–365.
17. Lambert D. Zero-inflated Poisson regression, with an application to defects in manufacturing. *Technometrics* 1992, 34(1):1–14.
18. Yau KKW, Wang K, Lee AH. Zero-Inflated Negative Binomial Mixed Regression Modeling of Over-Dispersed Count Data with Extra Zeros. *Biometrical Journal* 2003, 45(4):437–452.
19. Winkelmann R. *Econometric Analysis of Count Data*. 5th ed. Berlin: Springer-Verlag; 2008.
20. Wakefield J, Best N, Waller L. Bayesian approaches to disease mapping. In: Elliott P, Wakefield J, Best N, Briggs D, editors. *Spatial Epidemiology: Methods and Applications*. Oxford: Oxford University Press 2000. p. 106–127.
21. Gelman A, Carlin JB, Stern HS, Rubin DB. *Bayesian Data Analysis, Second Edition (Texts in Statistical Science)*. 2nd ed. Boca Raton, FL: Chapman & Hall/CRC Press, 2003.
22. Fernández C, Green PJ. Modelling spatially correlated data via mixtures: a Bayesian approach. *J Roy Stat Soc B* 2002, 64(4):805–826.
23. Agarwal DK, Gelfand AE, Citron-Pousty S. Zero-inflated models with application to spatial count data. *Environ Ecol Stat* 2002, 9(4):341–355.
24. Musal M, Aktekin T. Bayesian spatial modeling of HIV mortality via zero-inflated Poisson models. *Statistics in medicine* 2013, 32(2):267–81.
25. Lawson AB. *Bayesian Disease Mapping: Hierarchical Modeling in Spatial Epidemiology*. 1st ed. Boca Raton, FL: Chapman and Hall/CRC Press, 2008.
26. Martin TG, Wintle BA, Rhodes JR, Kuhnert PM, Field SA, et al. Zero tolerance ecology: improving ecological inference by modelling the source of zero observations. *Ecol Lett*. 2005, 8:1235–1246.
27. Duarte HHP, França EB. Qualidade dos dados da vigilância epidemiológica da dengue em Belo Horizonte, MG. *Rev Saúde Pública* 2006, 40:134-42.
28. Besag J, York J, Molli A. Bayesian image restoration, with two applications in spatial statistics. *Ann Inst Statist Math* 1991, 43(1):1–20.
29. McCullagh P, Nelder JA. *Generalized Linear Models*. 2nd ed. Boca Raton, FL: Chapman and Hall/ CRC Press, 1998.

30. Rue H, Martino S, Chopin N. Approximate Bayesian inference for latent Gaussian models by using integrated nested Laplace approximations. *J R Stat Soc Series B Stat Methodol* 2009, 71(2):319–392.
31. Li Y, Brown P, Rue H, al-Maini M, Fortin P. Spatial modelling of lupus incidence over 40 years with changes in census areas. *J Roy Stat Soc C-App* 2012, 61(1):99–115.
32. Rue H, Martino S. INLA: Functions which allow to perform a full Bayesian analysis of structured additive models using Integrated Nested Laplace Approximation. R package version 0.0. 2009.
33. Spiegelhalter DJ, Best NG, Carlin BP, Linde A Van Der. Bayesian measures of model complexity and fit. *J Roy Stat Soc B* 2002, 64(4):583–639.
34. Gneiting T, Raftery AE. Strictly Proper Scoring Rules, Prediction, and Estimation. *J Am Stat Assoc* 2007, 102(477):359–378.
35. Held L, Schrödle B, Rue H. Posterior and Cross-validatory Predictive Checks: A Comparison of MCMC and INLA. In: Kneib T, Tutz G, editors. *Statistical Modelling and Regression Structures*. Physica-Verlag HD 2010, 91–110.
36. Richardson S, Thomson A, Best N, Elliott P. Interpreting posterior relative risk estimates in disease-mapping studies. *Environ Health Perspect*. 2004, 112(9):1016–1025.
37. Malavasi HG. Análise espacial da epidemia de dengue em Campinas/ SP no ano de 2007.[Dissertação de Mestrado]. Campinas: Faculdade de Ciências Médicas, Universidade Estadual de Campinas; 2011.
38. Mondini A, Chiaravalloti Neto F: Socioeconomic variables and dengue transmission. *Cad Saúde Pública* 2007, 41:923-930.
39. Almeida ASD, Medronho RDA, Valencia LIO: Spatial analysis of dengue and the socioeconomic context of the city of Rio de Janeiro (Southeastern Brazil). *Rev Saúde Pública* 2009, 43:666-673.
40. Flauzino RF, Santos RS, Oliveira RM. Dengue, geoprocessamento e indicadores socioeconômicos e ambientais: um estudo de revisão, *Rev Panam Salud Publica*. 2009, 25(5):456–61.
41. Machado JP, Oliveira RMD, Souza-Santos R: Spatial analysis of dengue occurrence and living conditions in Nova Iguaçu, Rio de Janeiro State, Brazil. *Cad Saúde Pública* 2009, 25:1025-1034.
42. Pessanha JEM, Caiaffa WT, KrooLn EG, Proietti FA: Dengue em três distritos sanitários de Belo Horizonte, Brasil: inquérito soropidemiológico de base populacional, 2006 a 2007. *Rev Panam Salud Publica* 2010, 27:252-258.

43. Resendes APC, Silveira NAPR, Sabroza PC, Souza-Santos R: Determinação de áreas prioritárias para ações de controle da dengue. *Rev Saúde Pública* 2010, 44:274-282.
44. Teixeira TRA, Cruz OG: Spatial modeling of dengue and socio-environmental indicators in the city of Rio de Janeiro, Brazil. *Cad Saúde Pública* 2011, 27:591-602.
45. Rue H, Martino S. (2006). Approximate Bayesian inference for hierarchical Gaussian Markov random fields models. *J Statist Plann Inference* 2006, 137:3177–3192.
46. Openshaw S: Ecological fallacies and the analysis of areal census data. *Environ Plann A* 1984, 16:17-31.
47. Teixeira MG, Barreto ML, Costa MCN, Ferreira LDA, Vasconcelos P. Dinâmica de circulação do vírus da dengue em uma área metropolitana do Brasil. *Epidemiol Serv Saude* 2003, 12(2):87-97.
48. Barcellos C, Ramalho WM, Gracie R *et al.* Georreferenciamento de dados de saúde na escala submunicipal: algumas experiências no Brasil. *Epidemiol. Serv. Saúde* 2008, 17(1): 59-70.
49. Almeida MCM, Caiaffa WT, Assunção RM, Proietti FA. Dinâmica intra-urbana das epidemias de Dengue em Belo Horizonte, MG, 1996-2002. *Cad Saúde Pública* 2008; 24(10): 2385-2395.
50. Jackson C, Best N, Richardson S (2005) Improving ecological inference using individual-level data. *Stat Med* 25, 2136–2159.

5. DISCUSSÃO E CONCLUSÃO GERAL

A utilização de técnicas de geoprocessamento em pesquisas na área de saúde pública tem sido crescente nas últimas décadas. Informações de diferentes fontes e formatos podem ser integradas espacialmente, por meio de operações espaciais num ambiente de Sistemas de Informação Geográficas (SIG). Estudos pautados em indicadores socioeconômicos e ambientais vêm sendo utilizados na tentativa de compreensão da complexa relação de fatores que condicionam ou determinam a transmissão e manutenção da dengue.

Na análise da distribuição espacial de dados agregados, diferentes escalas geográficas para agregação das variáveis podem produzir variação nos resultados. Este problema, denominado problema da unidade de área modificável (*Modifiable Areal Unit Problem – MAUP*), pode ser agravado quando existem desigualdades socioeconômicas e de infraestrutura e, à medida que a agregação aumenta, a informação das covariáveis poderá discriminar menos as regiões e, conseqüentemente, os resultados poderão tornar-se pouco representativos do fenômeno em estudo. Sendo assim, estudar o efeito de diferentes formas de agregação dos dados no espaço e no tempo pode ser relevante para os estudos da distribuição dos riscos da doença.

A partir destas considerações, a presente tese analisou a distribuição espacial do risco de dengue e sua relação com fatores socioambientais, adotando-se o espaço geográfico intramunicipal de Campinas. Em escala local, a incidência da doença foi analisada por áreas de cobertura dos centros de saúde e por setores censitários. Ao longo do estudo foi enfatizado o aspecto metodológico na análise das taxas de risco da doença considerando as possibilidades de análises para os diferentes níveis de desagregação dos dados de dengue.

No primeiro artigo, realizou-se a análise da distribuição espacial do risco relativo de dengue relacionando-a com condições socioambientais, segundo as áreas de cobertura dos centros de saúde (ACCS). Modelos hierárquicos bayesianos espaciais foram ajustados, considerando-se a que a contagem de casos de dengue em cada ACCS seguia uma distribuição de Poisson. A variabilidade extra-Poisson decorrente da agregação espacial dos dados de contagem foi ajustada com a inclusão de efeitos aleatórios, modelando a dependência espacial por uma distribuição a *priori* autorregressiva condicional (CAR). O método INLA foi empregado para realizar as estimativas dos parâmetros.

Para os casos do grupo etário > 14 anos, os resultados evidenciaram o efeito das condições socioambientais sobre a distribuição espacial do risco de dengue, em que o risco da doença foi maior em áreas com condições mais precárias em relação a áreas com melhores condições socioambientais. Compreendendo a heterogeneidade espacial como um determinado espaço geográfico, em que se encontram populações em diferentes estratos de condições de vida, os resultados obtidos demonstraram que, para este grupo etário, a heterogeneidade no risco de dengue em Campinas é uma expressão da heterogeneidade nas condições socioambientais das ACCS. Para os casos do grupo etário menor ou igual a 14 anos esse efeito não foi observado.

A agregação dos casos por ACCS tem sua importância para a vigilância em saúde por expressar, neste território, a forma como estão organizadas as equipes envolvidas no controle da dengue e a atenção aos demais eventos de saúde como um todo. É também a unidade espacial adotada pela Secretária Municipal de Saúde (SMS) para produção e divulgação das estatísticas de saúde de dengue e demais agravos a saúde no município. Todavia, ao analisar os dados neste nível de agregação, alguns aspectos metodológicos apontam para uma maior reflexão sobre os resultados obtidos.

Uma característica atraente dos desenhos ecológicos na análise de dados em saúde é permitir relacionar informações socioeconômicas, demográficas e de saúde. Ressalta-se, portanto, a importância de fazer coincidir os limites dessas áreas com uma agregação de setores censitários, que permita a utilização das informações geradas no Censo. Para as ACCS, a delimitação dos polígonos tem por base os setores censitários. Contudo, esta agregação não é mutuamente exclusiva e alguns setores fazem parte de duas ACCS. Os dados do Censo 2000 disponibilizados pela SMS para as ACCS são, na verdade, estimativas realizadas pela SMS a partir das informações dos setores que compõem cada ACCS. Desta forma, a qualidade dos dados secundários fica comprometida e os resultados obtidos na modelagem dos dados podem ser viesados.

As ACCS são, em essência, uma reunião de áreas homogêneas e áreas heterogêneas em relação as suas condições de vida, que juntas moldam uma grande área heterogênea internamente. Se por um lado as ACCS têm importância como espaço geográfico integrador das ações de vigilância em saúde, por outro, em razão da extensão

das áreas, podem não refletir adequadamente o fenômeno em estudo, bem como, as estatísticas agregadas podem dissimular diferenças importantes.

É fundamental que as áreas utilizadas como unidade de análise sejam razoavelmente homogêneas quanto ao perfil da população residente. Entretanto, quanto maior o nível de homogeneidade interna da área, menor ela será, e conseqüentemente em muitas áreas não serão observadas contagens do evento de interesse.

O segundo artigo, buscou trabalhar os questionamentos apontados acima. As estimativas do risco relativo foram realizadas por setores censitários, sendo estes a menor unidade disponível para fins de agregação dos dados secundários. O modelo de Poisson não se ajusta a dados com muitos zeros, pois, este número elevado de zeros provoca uma dispersão elevada. Assim, para não agregar as áreas sem nenhuma contagem a outras com contagens, buscamos modelos alternativos ao de Poisson para modelar a superdispersão dos dados. Para tanto, foram utilizados os modelos de Poisson inflacionado de zero e o modelo binomial negativo inflacionado de zero.

Esses modelos são capazes de ajustar a variabilidade extra-Poisson e o excesso de zeros decorrente da agregação dos dados para pequenas áreas geográficas. Simultaneamente, efeitos aleatórios espaciais foram incluídos nos modelos de regressão. Foi analisado o efeito das condições socioambientais sobre a distribuição espacial do risco relativo de dengue. O modelo de Poisson inflacionado de zeros (ZIP) ajustou melhor os dados, evidenciado a importância de modelos alternativos em relação ao modelo padrão de Poisson na estimativa espacial do risco relativo.

O efeito das condições socioambientais não se fez presente por setor censitário, em oposição ao resultado encontrado no primeiro artigo por ACCS, indicando que as condições socioambientais não foram determinantes para heterogeneidade espacial do risco de dengue na cidade de Campinas, durante o período estudado. Desta forma, a variação espacial no risco relativo pode ser atribuída a variáveis não mensuradas de natureza espacial, a características locais de cada setor censitário. Também foi possível detectar aglomerados de setores com excesso de risco.

Nos estudos epidemiológicos que estudam a distribuição espacial de eventos de saúde, a presença de superdispersão e excesso de zeros é um assunto que tem sido pouco

abordado. Em termos da literatura nacional, a presente tese buscou preencher esta lacuna na nossa literatura.

Como possíveis trabalhos futuros, podemos explorar a aplicação do modelo de Poisson generalizado inflacionado de zeros (ZIGP), segundo os métodos aproximados de Laplace. Este método ainda não se encontra implementado no pacote R-INLA, o que torna um grande desafio esta possibilidade de análise. Podemos também analisar a distribuição espaço-temporal para pequenas áreas e analisar a independência espacial dos dados a partir de diferentes estruturas de vizinhança.

Portanto, os resultados apresentados nestes dois artigos evidenciaram: a relevância de se investigar diferentes escalas de agregação espacial, a necessidade de se aplicar modelos adequados para dados de pequenas áreas, a adequação da abordagem hierárquica bayesiana na redução de flutuações aleatórias nos riscos, a viabilidade de se obter estimativas rápidas e precisas com o método INLA, a importância da utilização de Sistemas de Informações Geográficas, permitindo a identificação de heterogeneidade espacial nas localidades estudadas e a construção de indicadores apropriados para estudos locais em função de suas particularidades.

6. REFERÊNCIAS BIBLIOGRÁFICAS

1. Tauil PL. Urbanization and dengue ecology. *Cad Saude Publica*. 2001;17 Suppl:99–102.
2. Guzmán MG, Kourí G. Dengue: an update. *Lancet Infect Dis*. Elsevier; 2002;2(1):33–42.
3. OMS. WHO. World Health Organization. Global strategy for dengue prevention and control 2010-2020 [Internet]. Global strategy for dengue prevention and control 2010-2020. 2012 [citado em: 30 nov . 2012]. Disponível em: http://apps.who.int/iris/bitstream/10665/75303/1/9789241504034_eng.pdf
4. BRASIL Ministério da Saúde. Fundação Nacional de Saúde. Dengue instruções para pessoal de combate ao vetor: manual de normas técnicas. 3rd ed. Brasília: Ministério da Saúde: Fundação Nacional de Saúde; 2001. p. 84.
5. BRASIL Ministério da Saúde. Secretaria de Vigilância em Saúde. Vigilância Epidemiológica. Vigilância Epidemiológica. Dengue: diagnóstico e manejo clínico. 2. ed Brasília. 2nd ed. Brasília: Ministério da Saúde; 2005.
6. Donalísio MR. O dengue no espaço habitado. São Paulo: Humanismo, Ciência e Tecnologia; 1999. p. 195.
7. Teixeira MG, Barreto ML, Guerra Z. Epidemiologia e medidas de prevenção do dengue. *Inf Epidemiol SUS*. 1999;8(4):5–33.
8. Gubler DJ. Emerging vector-borne Flavivirus diseases: are vaccines the solution? *Expert Rev Vaccines*. 2011;10:563–5.
9. Sabchareon A, Wallace D, Sirivichayakul C, Limkittikul K, Chanthavanich P, Suvannadabba S, et al. Protective efficacy of the recombinant, live-attenuated, CYD tetravalent dengue vaccine in Thai schoolchildren: a randomised, controlled phase 2b trial. *Lancet*. 2012;6736(12):1–9.
10. Medronho RA. Dengue e o ambiente urbano. *Revista Brasileira de Epidemiologia*. 2006;9(2):159–61.
11. Donalísio MR, Glasser CM. Vigilância entomológica e controle de vetores do dengue. *Revista Brasileira de Epidemiologia*. 2002 Dec;5(3):259–79.
12. Chiaravalloti Neto F, Fiorin AM, Conversani DT, Cesarino MB, Barbosa AAC, Dibo MR, et al. Controle do vetor do dengue e participação da comunidade em Catanduva, São Paulo, Brasil. *Cad Saude Publica*. 2003;19(6):1739–49.
13. Gómez Dantés H, Ramos Bonifaz B, Tapia Conyer R. The risk of dengue transmission; a space for stratification. *Salud pública Méx*. 1995;37(supl):88–97.

14. Eisen L, Lozano-Fuentes S. Use of Mapping and Spatial and Space-Time Modeling Approaches in Operational Control of *Aedes aegypti* and Dengue. Tatem AJ, editor. PLoS Negl Trop Dis. Public Library of Science; 2009;3(4):7.
15. OPS.Organização Panamericana de Saúde. Dengue y dengue hemorrágico en las Américas: guías para su prevención y control. Washington, DC; 1995.
16. Gubler D. The emergence of epidemic dengue fever and dengue hemorrhagic fever in the Americas: a case of failed public health policy. Revista Panamericana de Salud Pública. 2005;17(4):221–4.
17. Gubler DJ, Clark GG. Dengue/dengue hemorrhagic fever: the emergence of a global health problem. Emerg Infect Dis. 1995;1(2):55–7.
18. Gubler DJ. Dengue and dengue hemorrhagic fever: its history and resurgence as a global public health problem. In: Gubler DJ, Kuno G, editors. Dengue and Dengue Hemorrhagic Fever. CAB International Press; 1997. p. 1–22.
19. OPS.Organização Panamericana de Saúde. Resurgimiento del dengue en las Américas. Boletín Epidemiológico. 1997;18(2):16.
20. Pinheiro FP. Dengue in the Americas 1980-1987. Epidemiological Bulletin of PAHO. 1989;10(1):1–8.
21. Guzmán MG, García G, Kourí G. El dengue y el dengue hemorrágico: prioridades de investigación. Revista Panamericana de Salud Pública. 2006; 19(3):204–15.
22. Osanai CH, Rosa AP, Tang AT, Amaral RS do, Passos AD, Tauil PL. Surto de dengue em Boa Vista, Roraima. Nota prévia. Rev. Inst. Med. Trop. São Paulo. 1983;25(1):53–4.
23. Schatzmayr HG, Nogueira RMR, Rosa APAT da. An outbreak of dengue virus at Rio de Janeiro - 1986. Mem Inst Oswaldo Cruz. 1986;81(2):245–6.
24. Nogueira RMR, De Araújo JMG, Schatzmayr HG. Dengue viruses in Brazil, 1986-2006. Revista Panamericana de Salud Públicaalud Pública. 2007; 22(5):358–63.
25. Nogueira RMR, Miagostovich MP, Filippis AMB de, Pereira MAS, Schatzmayr HG. Dengue virus type 3 in Rio de Janeiro, Brazil. Mem Inst Oswaldo Cruz. 2001;96(7):925–6.
26. Siqueira-Júnior JB, Vinhal LC, Said RFC, Hoffmann, J L Martins J, Barbiratto SB. Dengue no Brasil: tendências e mudanças na epidemiologia, com ênfase nas epidemias de 2008 e 2010. In: Saúde M da, editor. Saúde Brasil 2010: uma análise da situação de saúde e de evidências selecionadas de impacto de ações de vigilância em saúde. Brasília; 2011. p. 157–71.

27. Temporão JG, Penna GO, Carmo EH, Coelho GE, Socorro Silva Azevedo R, Nunes MRT, et al. Dengue Virus Serotype 4, Roraima State, Brazil. *Emerg Infect Dis.* 2011;17(5):938–40.
28. Forshey B, Morrison AC, Cruz C, Rocha C, Vilcarrromero S, Guevara C, et al. Dengue Virus Serotype 4, Northeastern Peru, 2008. *Emerg Infect Dis.* 2009;15:1815–8.
29. Li D, Liu W, Guigon A, Mostyn C, Grant R, Aaskov J. Rapid Displacement of Dengue Virus Type 1 by Type 4, Pacific Region, 2007–2009. *Emerg Infect Dis.* 2010;16(1):123–5.
30. Rocco I. Ocorrência de Dengue no Estado de São Paulo, Brasil, de 1986 a 1996. *Revista do Instituto Adolfo Lutz.* 1998;
31. Spinola R. Dengue: velhos e novos desafios [Internet]. BEPA – Boletim Epidemiológico Paulista. 2004 [Acesso em: 20 dez. 2012]. p. 11–2. Disponível em: <http://www.cve.saude.sp.gov.br/agencia/bepa4.htm>
32. Centro de Vigilância Epidemiológica.Divisão de Zoonoses.Superintendência de Controle de Endemias.Secretaria de Estado da Saúde de São Paulo. Dengue no Estado de São Paulo [Internet]. BEPA – Boletim Epidemiológico Paulista. 2012 [Acesso em: 15 dez. 2012]. p. 22–40. Disponível em: http://www.cve.saude.sp.gov.br/bepa/txt/bepa98_dengue.htm
33. Centro de Vigilância Epidemiológica.Divisão de Zoonoses.Superintendência de Controle de Endemias.Secretaria de Estado da Saúde de São Paulo. Identificação do sorotipo DENV 4, GVE 29 – São José do Rio Preto, municípios de São José do Rio Preto e Paulo de Faria, SP, Brasil [Internet]. 2011 [Acesso em: 21 dez. 2012]. p. 25–8. Disponível em: http://www.cve.saude.sp.gov.br/bepa/txt/bepa89_dengue.htm
34. Rocco IM, Silveira VR, Maeda AY, Santos SJ, Silva, Suzuki, Carine Spennassatto Bisordi I, et al. First isolation of dengue 4 in the state of São Paulo, Brazil, 2011. *Revista do Instituto de Medicina Tropical de São Paulo.* 54(1):49–51.
35. Centro de Vigilância Epidemiológica.Divisão de Zoonoses.Superintendência de Controle de Endemias.Secretaria de Estado da Saúde de São Paulo. Dengue. Dados Estatísticos. Casos autóctones de dengue por DIR e Município no Estado de São Paulo 1995 a 2006 [Internet]. Dengue. Dados Estatísticos. Casos autóctones de dengue por DIR e Município no Estado de São Paulo 1995 a 2006. Disponível em: http://www.cve.saude.sp.gov.br/htm/zoo/dengue_dados.html
36. Coordenadoria de Vigilância em Saúde.Prefeitura Municipal de Campinas.Secretaria Municipal de Saúde. Informe Dengue: Situação atual e risco do próximo verão [Internet]. 2010. Disponível em:

<http://2009.campinas.sp.gov.br/saude/vigilancia/informes/InformeDengueOutubro2010.pdf>

37. Claro LBL, Tomassini HCB, Rosa MLG. Prevenção e controle do dengue: uma revisão de estudos sobre conhecimentos, crenças e práticas da população. *Cad Saude Publica*. 2004 Dec;20(6):1447–57.
38. Silva LJ. O conceito de espaço na epidemiologia das doenças infecciosas. *Cad Saude Publica*. 1997 Oct;13(4):585–93.
39. Elliott P, Wakefield J, Best N, Briggs D. *Spatial Epidemiology: Methods and Applications*. London: Oxford University Press; 2001. p. 475.
40. Sabroza PC, Kawa H, Campos WSQ. Doenças transmissíveis: ainda um desafio. In: Minayo MCS, editor. *Os Muitos Brasis: Saúde e População na Década de 80*. São Paulo: Hucitec; 1995. p. 177–244.
41. Barcellos C, Bastos FI. Geoprocessamento, ambiente e saúde: uma união possível? *Cad Saude Publica*. 1996 Sep;12(3):389–97.
42. Barata RB. *Condições de vida e situação de saúde*. Barata RB, editor. Rio de Janeiro: ABRASCO; 1997.
43. Flauzino RF, Souza-Santos R, Barcellos C, Gracie R, Magalhães MAFM, Oliveira RM. Spatial heterogeneity of dengue fever in local studies, City of Niterói, Southeastern Brazil. *Cad Saude Publica*. 2009;43(6):1035–43.
44. Susser M. The logic in ecological: II. The logic of design. *AmJPublic Health*. 1994;84(5):830–5.
45. Lawson AB, Biggeri A, Böhning D, Lesaffre E, Viel JF, Bertollini R. Disease mapping and risk assessment for public health. *Public Health*. Wiley; 1999. p. 502 p.
46. Flauzino RF, Souza-Santos R, Barcellos C, Gracie R, Magalhães MAFM, Oliveira RM. Heterogeneidade espacial da dengue em estudos locais , Niterói , RJ. *Cad Saude Publica*. 2009;43(6):1035–43.
47. Rojas LI, Barcellos C, Peiter P. Utilização de mapas no campo da Epidemiologia no Brasil: reflexões sobre trabalhos apresentados no IV Congresso Brasileiro de Epidemiologia. *Informe Epidemiológico do Sus*. 1999 ;8(2):27–35.
48. Druck S, Carvalho MS, Câmara G, Monteiro AVM, editors. *Análise Espacial de Dados Geográficos*. Brasília: EMBRAPA; 2004.
49. Bailey TC. *Spatial statistical methods in health*. *Cad Saude Publica*. 2001 Oct;17(5):1083–98.
50. Lawson AB. *Bayesian Disease Mapping: Hierarchical Modeling in Spatial Epidemiology*. 1st ed. Boca Raton, FL: Chapman and Hall/CRC Press; 2008. p. 368.

51. Wakefield J, Best N, Waller L. Bayesian approaches to disease mapping. In: Elliott P, Wakefield J, Best N, Briggs D, editors. *Spatial Epidemiology: Methods and Applications*. Oxford: Oxford University Press; 2000. p. 106–27.
52. Wakefield J. Disease mapping and spatial regression with count data. *Biostatistics* (Oxford, England). 2007 Apr 1;8(2):158–83.
53. Gelman A, Carlin JB, Stern HS, Rubin DB. *Bayesian Data Analysis, Second Edition* (Texts in Statistical Science). 2nd ed. Book. Boca Raton, FL: Chapman & Hall/CRC Press; 2003. p. 668.
54. Mondini A, Chiaravalloti Neto F. Socioeconomic variables and dengue transmission. *Cad Saude Publica*. 2007;41(6):923–30.
55. Resendes APC, Silveira NAPR, Sabroza PC, Souza-Santos R. Determinação de áreas prioritárias para ações de controle da dengue. *Rev Saude Publica*. 2010;44(2):274–82.
56. Hair JF, Black B, Babin B, Anderson RE, Tatham RL. *Multivariate Data Analysis*. 6th ed. Prentice Hall; 2005. p. 928.
57. McCullagh P, Nelder JA. *Generalized Linear Models*. 2nd ed. *Generalized Linear Models Second Edition*. Boca Raton, FL: Chapman and Hall/ CRC Press; 1998. p. 511.
58. Elliott P, Wartenberg D. *Spatial Epidemiology: Current Approaches and Future Challenges*. *Environ Health Perspect*. National Institute of Environmental Health Sciences; 2004;112(9):998–1006.
59. Clayton D, Kaldor J. Empirical Bayes estimates of age-standardized relative risks for use in disease mapping. *Biometrics*. JSTOR; 1987;43(3):671–81.
60. Lambert D. Zero-inflated Poisson regression, with an application to defects in manufacturing. *Technometrics*. JSTOR; 1992;34(1):1–14.
61. Martin TG, Wintle BA, Rhodes JR, Kuhnert PM, Field SA, Low-Choy SJ, et al. Zero tolerance ecology: improving ecological inference by modelling the source of zero observations. *Ecology Letters*. Wiley Online Library; 2005;8(11):1235–46.
62. Cameron AC, Trivedi PK. *Regression Analysis of Count Data*. Hammond P, Holly A, editors. Society. Cambridge University Press; 1998. p. 411.
63. Yau KKW, Wang K, Lee AH. Zero-Inflated Negative Binomial Mixed Regression Modeling of Over-Dispersed Count Data with Extra Zeros. *Biometrical Journal*. Wiley Online Library; 2003;45(4):437–52.
64. Duarte H, França E. Qualidade dos dados da vigilância epidemiológica da dengue em Belo Horizonte, MG. *Rev Saúde Pública*. 2006;(40):134–42.

65. Besag J, York J, Molli A. Bayesian image restoration, with two applications in spatial statistics. *Ann Inst Statist Math*. Springer Netherlands; 1991;43(1):1–20.
66. Fahrmeir L, Osuna L. Structured additive regression for overdispersed and zero-inflated count data. *Applied Stochastic Models in Business and Industry*. Wiley Online Library; 2006;22: 351–69.
67. Kelsall J, Wakefield J. Discussion of “Bayesian models for spatially correlated disease and exposure data”. In: Bernardo J, Berger J, Dawid A, Smith A, editors. *Bayesian Statistics 6*. Oxford, UK: Oxford University Press; 1999. p. 131–56.
68. Knorr-Held L, Rue H. On Block Updating in Markov Random Field Models for Disease Mapping. *Scandinavian Journal of Statistics*. Blackwell Publishers; 2002;29(4):597–614.
69. Rue H, Follstad T. Gaussian Markov Random Field Models With Applications in Spatial Statistics. *Science And Technology*. 2003;(2002):1–20.
70. Rue H, Martino S, Chopin N. Approximate Bayesian inference for latent Gaussian models by using integrated nested Laplace approximations. *J R Stat Soc Series B Stat Methodol*. Wiley Online Library; 2009;71(2):319–92.
71. Geisser S. *Predictive Inference (Chapman & Hall/CRC Monographs on Statistics & Applied Probability)*. Chapman and Hall/CRC; 1993. p. 240.

7. Anexos

7.1. Anexo 1. Comprovante de aceite do artigo submetido ao Cad. Saúde Pública

CSP

CADERNOS DE SAÚDE PÚBLICA REPORTS IN PUBLIC HEALTH

SECRETARIA DE
CADERNOS DE SAÚDE
PÚBLICA
ESCOLA NACIONAL DE SAÚDE
PÚBLICA SERGIO AROUCA
FUNDAÇÃO OSWALDO CRUZ
Rua Leopoldo Bulhões 1480
21041-210 Rio de Janeiro, RJ,
Brasil
Telefone: (+55-21) 2598-2511
Telefax: (+55-21) 2598-2737
cadernos@ensp.fiocruz.br
http://www.enp.fiocruz.br/csp


Rio de Janeiro, 31 de janeiro de 2013.

Ilmo Sr.
Dr. José Vilton Costa

Em nome do Conselho Editorial de **Cadernos de Saúde Pública**, comunicamos que o artigo de sua autoria, em colaboração com Maria Rita Donalisio Cordeiro & Liciane Vaz de Arruda Silveira, intitulado "*Distribuição Espacial da Incidência de Dengue e Condições Socioambientais em Campinas, Estado de São Paulo, Brasil, 2007*" foi aprovado quanto ao seu mérito científico.

A conclusão do processo editorial de seu artigo dependerá da avaliação técnico-editorial com vistas a detectar dúvidas de formatação, referências bibliográficas, figuras e/ou tabelas. Comunicação nesse sentido lhe será enviada oportunamente.

Atenciosamente,


EDITORAS
EDITORS
Marília Sá Carvalho
Claudia Travassos
Claudia Medina Coeli


EDITORES ASSOCIADOS
ASSOCIATE EDITORS
Francisco I. Bastos
Paulo Marchiori Buss
Maria Cristina Marino Calvo
Suely F. Deslandes
Gilberto Kac
Ronir Raggio Luiz
Mônica Silva Martins
Paulo R. Menezes
Maria Teresa Anselmo Olinto
Claudia Garcia Serpa Osorio-de-
Castro
Michael Reichenheim
Vilma Sousa Santana
Iná S. Santos
Ricardo Ventura Santos
Antônio Augusto Moura da Silva
Reinaldo Souza-Santos
Ligia Maria Vieira da Silva
Wayner Vieira de Souza
Guilherme Loureiro Werneck


EDITOR DE ARTIGOS DE REVISÃO
REVIEW EDITOR
Edison Iglesias de Oliveira Vidal

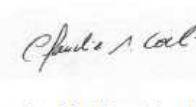
EDITORA DE RESENHAS
BOOK REVIEW EDITOR
Martha Cristina Nunes Moreira

EDITORES ASSISTENTES
ASSISTANT EDITORS
Leandro Carvalho
Marcia Pietrukowicz
Carolina Ribeiro

EDITORA ADMINISTRATIVA
ADMINISTRATIVE EDITOR
Carla Alves


Marília Sá Carvalho
Editora


Claudia Travassos
Editora


Claudia Medina Coeli
Editora