

INDICADOR DE RISCO DE ÓBITO DE PACIENTES COM AIDS NO  
MUNICÍPIO DE CAMPINAS SOB UMA ABORDAGEM DE  
MODELOS ESPAÇO-TEMPORAIS EM ANÁLISE DE  
SOBREVIVÊNCIA

**Thiago Santos Mota**

Tese apresentada à Universidade Estadual Paulista “Júlio de Mesquita Filho” para a obtenção do título de Doutor em Biometria.

BOTUCATU  
São Paulo - Brasil  
Junho – 2018

INDICADOR DE RISCO DE ÓBITO DE PACIENTES COM AIDS NO  
MUNICÍPIO DE CAMPINAS SOB UMA ABORDAGEM DE  
MODELOS ESPAÇO-TEMPORAIS EM ANÁLISE DE  
SOBREVIVÊNCIA

**Thiago Santos Mota**

Orientadora: Profa. Dra. **Liciana Vaz de Arruda Silveira**

Tese apresentada à Universidade Estadual Paulista “Júlio de Mesquita Filho” para a obtenção do título de Doutor em Biometria.

BOTUCATU  
São Paulo - Brasil  
Junho – 2018

## Ficha Catalográfica

FICHA CATALOGRÁFICA ELABORADA PELA SEÇÃO TÉC. AQUIS. TRATAMENTO DA INFORM.  
DIVISÃO TÉCNICA DE BIBLIOTECA E DOCUMENTAÇÃO - CÂMPUS DE BOTUCATU - UNESP  
BIBLIOTECÁRIA RESPONSÁVEL: ROSANGELA APARECIDA LOBO-CRB 8/7500

Mota, Thiago Santos.

Indicador de risco de óbito de pacientes com AIDS no município de Campinas sob uma abordagem de modelos espaço-temporais em análise de sobrevivência / Thiago Santos Mota. - Botucatu, 2018

Tese (doutorado) - Universidade Estadual Paulista "Júlio de Mesquita Filho", Instituto de Biociências de Botucatu

Orientador: Liciane Vaz de Arruda Silveira  
Capes: 90194000

1. AIDS (Doença). 2. Mortalidade. 3. Terapia antirretroviral de alta atividade. 4. Análise espacial (Estatística). 5. Análise de sobrevivência (Biometria).

Palavras-chave: AIDS; Distribuição espacial; HAART; INLA; Mortalidade.

## **Dedicatória**

Dedico este trabalho aos meus pais, Severina e Antonio (in memoriam),  
que proporcionaram uma infância alegre e com muitos ensinamentos.

## Agradecimentos

Primeiramente, agradeço aos meus pais, Severina e Antonio (in memoriam), pela minha educação, sem eles não seria a pessoa que sou hoje.

A minha esposa, Flávia Priscila Ventura, companheira de todas as horas que compartilhou todos momentos da elaboração desta tese no dia a dia, as privações, alegrias e angústias.

A todos os professores do curso de doutorado em Biometria e demais funcionários do departamento de Bioestatística, que direta ou indiretamente contribuíram nesta etapa da minha formação.

À banca de qualificação, composta pelos professores Dr. José Eduardo Corrente e Dr. Carlos Magno Castelo Branco Fortaleza, pelas correções e sugestões que foram essenciais para a melhoria deste trabalho.

À profa. Dra. Licianá Vaz de Arruda Silveira pela atenciosa orientação na elaboração deste trabalho, colaborando assim na minha formação acadêmica.

À profa. Dra. Maria Rita Donalisio da Faculdade de Ciências Médicas da UNICAMP, pela colaboração e coorientação na discussão dos resultados, pelo suporte em sanar dúvidas do banco de dados e questões referentes a epidemia da AIDS.

À Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES), pela minha bolsa de doutorado no primeiro ano do curso.

# Sumário

	Página
<b>RESUMO</b>	<b>vii</b>
<b>SUMMARY</b>	<b>ix</b>
<b>1 INTRODUÇÃO</b>	<b>1</b>
<b>2 REVISÃO DE LITERATURA</b>	<b>3</b>
2.1 Análise de Sobrevivência . . . . .	3
2.1.1 Modelo de Cox . . . . .	4
2.1.2 Modelo de Cox estratificado . . . . .	7
2.2 Modelo Linear Generalizado . . . . .	8
2.3 Modelo Aditivo Generalizado . . . . .	9
2.4 Inferência Bayesiana . . . . .	12
2.5 Método de estimação INLA . . . . .	15
2.6 Modelo espaço-temporal na análise de dados de Sobrevida . . . . .	16
2.7 AIDS . . . . .	18
<b>3 MATERIAL E MÉTODOS</b>	<b>21</b>
3.1 Dados . . . . .	21
<b>Artigo 1</b>	<b>23</b>
<b>Artigo 2</b>	<b>43</b>
<b>4 DISCUSSÃO E CONCLUSÃO GERAL</b>	<b>64</b>

**REFERÊNCIAS BIBLIOGRÁFICAS**

vi

**66**

**APÊNDICES**

**70**

**INDICADOR DE RISCO DE ÓBITO DE PACIENTES COM AIDS NO  
MUNICÍPIO DE CAMPINAS SOB UMA ABORDAGEM DE  
MODELOS ESPAÇO-TEMPORAIS EM ANÁLISE DE  
SOBREVIVÊNCIA**

Autor: THIAGO SANTOS MOTA

Orientadora: Prof. Dr. LICIANA VAZ DE ARRUDA SILVEIRA

**RESUMO**

O objetivo principal desta tese foi avaliar a sobrevida e o risco espacial de óbito de pacientes portadores de HIV/AIDS em três períodos de tempo, por meio de modelos de sobrevida espaço-temporais. A justificativa desse estudo se deve a lacuna de trabalhos na área de epidemiologia que abordem simultaneamente técnicas estatísticas de análise espacial e sobrevivência na análise de dados de AIDS. Essas técnicas são úteis para um melhor entendimento sobre a epidemia, para auxiliar no monitoramento clínico, avaliar a mortalidade e obter os fatores de risco que influenciam na sobrevida de pacientes portadores de HIV/AIDS. Os dados foram obtidos do Sistema de Informação de Agravos de Notificação (SINAN) e dos óbitos que constam no banco de dados do sistema de informação sobre mortalidade (SIM) de moradores de Campinas. O banco dados foi dividido em três coortes retrospectivas, sendo a primeira composta de 286 indivíduos notificados nos primeiros anos



da epidemia de 1980 a 1990 (coorte 1), uma segunda coorte com 1456 indivíduos notificados de 1996 a 2000 e uma terceira coorte com 1342 indivíduos notificados de 2001 a 2005 (coorte 3). Nestas coortes o tempo de seguimento médio foi de 10 anos. Na modelagem utilizou-se duas abordagens, na primeira, bayesiana, ajustou-se um modelo semi-paramétrico bayesiano empregando o método de Aproximação de Laplace Aninhada e Integrada (INLA) agregando as localizações das residências dos indivíduos que pertenciam ao mesmo setor censitário. Na segunda abordagem, frequentista, utilizou-se uma extensão do modelo de risco proporcionais, incorporado nos Modelos Lineares Aditivos Generalizados (GAM), considerando uma função de suavização com as localizações espaciais dos indivíduos (latitude e longitude). Como resultado, obteve-se os mapas de risco que apontaram tendência de queda da mortalidade e evidenciaram os setores que apresentaram maior e menor risco, além disso obteve-se os fatores de risco de óbito dos indivíduos portadores de HIV/AIDS. As duas abordagens diferem em alguns resultados, mas indicam a diminuição do risco de morte por AIDS nos últimos anos, após o acesso à terapia anti-retroviral altamente ativa. Os mapas e os fatores de risco obtidos orientam possíveis ações e monitoramento da doença no município.

# INDICATOR OF DEATH RISK IN AIDS PATIENTS IN THE CITY OF CAMPINAS IN A SPATIO-TEMPORAL MODELS APPROACH TO SURVIVAL ANALYSIS

Author: THIAGO SANTOS MOTA

Adviser: Prof. Dr. LICIANA VAZ DE ARRUDA SILVEIRA

## SUMMARY

The main objective of this thesis was to evaluate the survival and spatial risk of death of patients with HIV/AIDS in three periods of time, through of spatio-temporal survival models. The justification for this study is due to a gap in work in the area of the epidemiology that simultaneously addresses statistical techniques of spatial analysis and survival in the analysis of AIDS data. These techniques are useful for a better understanding of the epidemic, to assist in clinical monitoring, to evaluate mortality and to obtain the risk factors that influence the survival of patients with AIDS. The data were obtained from the Information System for Notifiable Diseases (SINAN) and from the deaths in the database of the mortality information system (SIM) of residents of Campinas. The data was divided into three retrospective cohorts, the first consisting of 286 individuals reported in the first years

of the epidemic from 1980 to 1990 (cohort 1), a second cohort with 1456 individuals reported from 1996 to 2000 and a third cohort with 1342 individuals reported from 2001 to 2005 (cohort 3). In these cohorts the mean follow-up time was 10 years. In the modeling we used two approaches, in the first, Bayesian, adjusted a Bayesian semi-parametric model using the Approximation Laplace Nested and Integrated (INLA) aggregating the locations of the residences of individuals belonging to the same census sector. In the second approach, frequentist, we used a proportional extension of the risk model, incorporated Generalized Additive Models (GAM), considering a spatial locations of individuals (latitude and longitude). As a result, we obtained the risk maps that indicated a tendency of mortality decrease and evidenced the sectors that presented higher and lower risk, in addition to the risk factors were obtained from individuals with HIV/AIDS. The two approaches differ in some indicate a reduction in the risk of death from AIDS in recent years, after access to highly active antiretroviral therapy. The maps and the risk factors obtained guide possible actions and monitoring of the disease in the municipality.

# 1 INTRODUÇÃO

Os primeiros casos da síndrome de Imunodeficiência Humana Adquirida – AIDS foram descritos nos EUA em 1981 (Gottlieb et al., 1981).

Embora, inicialmente, a categoria de transmissão fosse vinculada a homossexuais, o HIV/AIDS se disseminou rapidamente a outros seguimentos da sociedade (Fonseca & Bastos, 2007).

Avanços desde a década de 1990 no aperfeiçoamento à assistência ao paciente e à utilização da Terapia Antiretroviral Altamente Ativa (HAART), tiveram um grande impacto na redução do número de óbitos por AIDS (Killian & Levy, 2011).

O documento 90/90/90 proposto pela UNAIDS estabelece metas otimizadas de fim da epidemia, que preconiza 90% de diagnóstico, de tratamento e supressão viral em 2020 (UNAIDS, 2014). Estes percentuais continuam sendo um desafio para o Brasil e para o mundo, dadas as discrepâncias do acesso clínico associado às desigualdades sociais (UNAIDS, 2014; McMahon et al., 2011; Guibu et al., 2011).

Neste contexto, a informação epidemiológica adquiriu um papel principal na definição e condução de ações da vigilância da AIDS (Sousa & Pinto Júnior, 2016). A utilização da estatística espacial e análise de sobrevivência podem contribuir neste cenário auxiliando o gestor de saúde na tomada de decisão.

A estatística espacial pode proporcionar um melhor entendimento sobre a epidemia, auxiliar no monitoramento, contribuir para um melhor conhecimento dos riscos de transmissão da doença e fornecer uma visão geral dos determinantes sociais do processo saúde-doença, ainda incipientes na análise de indicadores de saúde (Sousa & Pinto Júnior, 2016).

A análise de sobrevivência pode auxiliar em estudos que avaliem o

impacto da mortalidade associada a AIDS, estudos de novas terapias e encontrar os fatores de risco que influenciam na sobrevivência de pacientes portadores de AIDS (Chequer et al., 1992).

Nesta tese o principal objetivo foi reunir estas duas ferramentas, estatística espacial e análise de sobrevivência, com o intuito de avaliar a sobrevida e o risco espacial de óbito em três períodos de tempo por meio do ajuste de modelos de sobrevida espaço-temporais, obtendo mapas que servem como indicadores de risco de óbito por AIDS na cidade de Campinas, São Paulo, nos períodos de: 1980–1990, 1996–2000 e 2001–2005.

## 2 REVISÃO DE LITERATURA

### 2.1 Análise de Sobrevivência

Análise de sobrevivência tem como objetivo explorar dados cuja variável resposta é o tempo até a ocorrência de um evento de interesse, o qual é denominado muitas vezes como falha. Aplicações desta análise estatística vêm sendo empregadas em diversas áreas como: medicina, engenharia, ciências sociais, financeiras, dentre outras. Vários problemas surgem nestas áreas como, por exemplo, o tempo até a morte de um paciente, o tempo até a falha de um componente eletrônico, entre outros. Esses dados são frequentemente censurados, ou seja, apresentam observações incompletas, que por algum motivo não foi possível observar a ocorrência do evento de interesse. Duas razões justificam considerar a censura na modelagem: (i) mesmo sendo incompletas, as observações censuradas fornecem informações sobre o tempo de vida de pacientes; (ii) a omissão das censuras no cálculo das estatísticas de interesse podem acarretar conclusões viciadas (Colosimo & Giolo, 2006).

Em análise de sobrevivência o tempo de falha é especificado pela função de sobrevivência ou pela função taxa de falha (ou risco).

A função de sobrevivência é a probabilidade de o indivíduo não falhar até o tempo  $T$ , ou seja, do indivíduo sobreviver ao tempo  $t$ , podendo ser escrita como:

$$S(t) = P(T \geq t) = 1 - F(t) = \int_t^{\infty} f(u) du, \quad t \geq 0,$$

observe que  $S(t)$  é uma função contínua e monótona decrescente com  $S(0) = 1$  e  $S(\infty) = \lim_{t \rightarrow \infty} S(t) = 0$ . Ocasionalmente, tem-se que  $S(\infty) > 0$  em situações em que é considerado que alguns indivíduos nunca falham.

A função taxa de falha ou risco  $\lambda(t)$  é definida como:

$$\lambda(t) = \lim_{\Delta t \rightarrow 0} \frac{P(t \leq T \leq t + \Delta t | T > t)}{\Delta t} = \frac{f(t)}{S(t)} = -\frac{d}{dt} \log \{S(t)\},$$

a função de risco específica a taxa instantânea de falha no tempo  $t$ , dado que o indivíduo sobreviveu até o tempo  $t$ . Segundo Colosimo & Giolo (2006), esta função é bastante útil para descrever a distribuição do tempo de vida de pacientes.

Outra função importante em análise de sobrevivência é a função de risco acumulado, pois fornece o risco acumulado do indivíduo até um tempo  $t$ . De acordo com Colosimo & Giolo (2006) esta função não tem interpretação direta, mas é bastante útil para avaliar a função de risco, principalmente no processo de estimação não paramétrica em que a função risco acumulado apresenta propriedades ótimas e a função de risco é difícil de ser estimada. A função de risco acumulado é definida como:

$$\Lambda(t) = \int_0^t \lambda(u) du = -\log \{S(t)\}.$$

### 2.1.1 Modelo de Cox

Uma das maneiras para modelar dados de sobrevivência na presença de covariáveis é utilizar o modelo de regressão semi-paramétrico proposto por Cox (1972), que também é denominado modelo de riscos proporcionais, pois a razão da taxa de falha de dois indivíduos distintos é constante no tempo. Assumindo que os tempos  $t_i$  sejam independentes, com  $i = 1, \dots, n$ , o modelo de Cox para o  $i$ -ésimo indivíduo, dado o vetor  $\mathbf{x} = (x_1, \dots, x_p)'$  de covariáveis, é dado por:

$$\lambda(t|\mathbf{x}_i) = \lambda_0(t) \exp \{\mathbf{x}_i' \boldsymbol{\beta}\}, \quad (1)$$

em que  $\lambda(t|\mathbf{x}_i)$  representa a função taxa de risco,  $\lambda_0(t)$  é o componente não paramétrico, conhecido como função de base ou basal, pois  $\lambda(t) = \lambda_0(t)$  para  $\mathbf{x} = 0$ ,  $\boldsymbol{\beta} = (\beta_1, \dots, \beta_p)'$  é o vetor dimensão  $1 \times p$  de parâmetros associados às covariáveis e  $\mathbf{x}_i$  é o vetor de dimensão  $p \times 1$  de covariáveis observadas para o  $i$ -ésimo indivíduo. A

representação da razão da taxa de falha de dois indivíduos distintos  $r$  e  $j$  é dada por:

$$\frac{\lambda_r(t)}{\lambda_j(t)} = \frac{\lambda_0 \exp \{ \mathbf{x}'_r \boldsymbol{\beta} \}}{\lambda_0 \exp \{ \mathbf{x}'_j \boldsymbol{\beta} \}} = \exp \{ \mathbf{x}'_r \boldsymbol{\beta} - \mathbf{x}'_j \boldsymbol{\beta} \}.$$

Para estimar o parâmetro  $\boldsymbol{\beta}$ , Cox (1972) propôs uma função de verossimilhança e no artigo seguinte, Cox (1975), ele introduziu o conceito de verossimilhança parcial obtendo a mesma função de verossimilhança do artigo Cox (1972) e estudou as suas propriedades. Suponha que uma amostra de censura aleatória  $(t_i, \delta_i)$ ,  $i = 1, \dots, n$ , existam  $k \leq n$  falhas distintas nos tempos  $t_1 < \dots < t_k$ . Seja  $R_i = R(t_i)$  denotando o conjunto de indivíduos que estão sob risco no tempo  $t_i$ . Colosimo & Giolo (2006) mostram uma forma simples de entender a função de verossimilhança parcial sem utilizar um processo de contagem, assim, considerando o seguinte argumento condicional: a probabilidade da  $i$ -ésima observação vir a falhar no tempo  $t_i$  conhecendo quais observações estão sob risco em  $t_i$ , então, tem-se:

$$\begin{aligned} \frac{P(\text{indivíduo falhar em } t_i | \text{sobreviveu a } t_i \text{ e história até } t_i)}{P(\text{uma falha em } t_i | \text{história até } t_i)} &= \\ \frac{\lambda_i(t_i | \mathbf{x}_i)}{\sum_{j \in R(t_i)} \lambda_j(t_i | \mathbf{x}_j)} &= \frac{\exp \{ \mathbf{x}'_i \boldsymbol{\beta} \}}{\sum_{j \in R(t_i)} \exp \{ \mathbf{x}'_j \boldsymbol{\beta} \}}, \end{aligned} \quad (2)$$

fazendo o produto de todos os termos representados na equação (2) associados aos tempos distintos de falha, obtém-se a seguinte função de verossimilhança:

$$L(\boldsymbol{\beta}) = \prod_{i=1}^n \left( \frac{\exp \{ \mathbf{x}'_i \boldsymbol{\beta} \}}{\sum_{j=1}^n Y_j(t_i) \exp \{ \mathbf{x}'_j \boldsymbol{\beta} \}} \right)^{\delta_i}, \quad (3)$$

em que a função indicadora de risco  $Y_j(t_i) = I(T_j \geq t_i)$ ,  $Y_j(t_i) = 1$  se e somente se  $j \in R(t_i)$ , e  $\delta_i$  é uma função indicadora de falha. Aplicando o logaritmo na equação (3), obtém-se a seguinte equação:

$$l(\boldsymbol{\beta}) = \sum_{i=1}^n \delta_i \left[ \mathbf{x}'_i \boldsymbol{\beta} - \log \left\{ \sum_{j=1}^n Y_j(t_i) \exp \{ \mathbf{x}'_j \boldsymbol{\beta} \} \right\} \right]. \quad (4)$$

Os valores de  $\boldsymbol{\beta}$  que maximizam a função de verossimilhança parcial,  $l(\boldsymbol{\beta})$ , são obtidos resolvendo o sistema de equações definido por



$U(\boldsymbol{\beta}) = \left( \frac{\partial l}{\partial \beta_1}, \dots, \frac{\partial l}{\partial \beta_p} \right)' = 0$ . Definindo para  $t > 0$  o vetor de dimensão  $p \times 1$ , dado por:

$$\bar{\mathbf{x}}(t, \boldsymbol{\beta}) = \frac{\sum_{j=1}^n Y_j(t) \mathbf{x}_j \exp \{ \mathbf{x}_j' \boldsymbol{\beta} \}}{\sum_{j=1}^n Y_j(t) \exp \{ \mathbf{x}_j' \boldsymbol{\beta} \}}, \quad (5)$$

em que  $\bar{\mathbf{x}}(t, \boldsymbol{\beta})$  é uma média ponderada do vetor de covariáveis dos indivíduos em risco no tempo  $t$ .

O vetor escore de derivadas  $U(\boldsymbol{\beta})$ , pode ser representado como:

$$U(\boldsymbol{\beta}) = \sum_{i=1}^n \delta_i [\mathbf{x}_i - \bar{\mathbf{x}}(t_i, \boldsymbol{\beta})]. \quad (6)$$

A matriz de informação  $\mathcal{I}(\boldsymbol{\beta}) = -\frac{\partial^2 l}{\partial \boldsymbol{\beta} \partial \boldsymbol{\beta}'}$  de dimensão  $p \times p$  pode ser vista como:

$$\mathcal{I}(\boldsymbol{\beta}) = \sum_{i=1}^n \delta_i \left\{ \frac{\sum_{j=1}^n Y_j(t_i) \exp \{ \mathbf{x}_j' \boldsymbol{\beta} \} [\mathbf{x}_j - \bar{\mathbf{x}}(t_i, \boldsymbol{\beta})] [\mathbf{x}_j - \bar{\mathbf{x}}(t_i, \boldsymbol{\beta})]'}{\sum_{j=1}^n Y_j(t_i) \exp \{ \mathbf{x}_j' \boldsymbol{\beta} \}} \right\}. \quad (7)$$

Segundo Lawless (2002), as equações de verossimilhança,  $U(\boldsymbol{\beta}) = 0$ , podem ser resolvidas utilizando o método iterativo de Newton-Raphson ou outros métodos, sendo que numerosos pacotes dão a estimativa  $\hat{\boldsymbol{\beta}}$  e erro padrão, testes ou intervalos de confiança baseados na aproximação assintótica normal,

$$\hat{\boldsymbol{\beta}} \simeq N_p \left( \boldsymbol{\beta}, [\mathcal{I}(\hat{\boldsymbol{\beta}})]^{-1} \right).$$

Um problema que surge neste processo de estimação é que a função de verossimilhança parcial, dada pela equação (3), assume que o tempo de sobrevivência é contínuo e com isso não há possibilidade de empates dos tempos de sobrevivência observados. Na prática, empates podem ocorrer nos tempos de falha ou de censura devido à escala de medida, ou seja, se for utilizado uma escala discreta (horas, dias, semanas, meses ou anos), e assim é frequente a observação de tempos de sobrevivência iguais para dois ou mais indivíduos no estudo. Para contornar este problema, Breslow (1972) e Peto (1972) propuseram uma aproximação que é frequentemente usada em programas estatísticos. Assim, seja  $\mathbf{s}_i$  o vetor formado pela soma das correspondentes  $p$  covariáveis para os indivíduos que falham no mesmo tempo  $t_i$  ( $i=1, \dots, n$ ) e  $d_i$  o

número de falhas neste mesmo tempo. A aproximação da equação (3) é escrita como:

$$L(\boldsymbol{\beta}) = \prod_{i=1}^k \frac{\exp \{s'_i \boldsymbol{\beta}\}}{\left[ \sum_{j \in R(t_i)} \exp \{ \mathbf{x}'_j \boldsymbol{\beta} \} \right]^{d_i}}, \quad (8)$$

sendo esta aproximação adequada quando o número de observações empatadas em qualquer tempo não é muito grande.

### 2.1.2 Modelo de Cox estratificado

Uma das pré-suposições do modelo de Cox é a suposição de riscos proporcionais. Uma forma de verificar esta hipótese após o ajuste do modelo é através do resíduo padronizado de Schoenfeld (1982). Para definir os resíduos no modelo de Cox, considere que o  $i$ -ésimo indivíduo é observado a falhar e possui como vetor de covariáveis  $\mathbf{x}_i = (x_{i1}, x_{i2}, \dots, x_{ip})$ , em que cada componente  $r_{iq}$ , para  $q = 1, \dots, p$ , é definido por:

$$r_{iq} = x_{iq} - \frac{\sum_{j \in R(t_i)} x_{jq} \exp \{ \mathbf{x}'_j \hat{\boldsymbol{\beta}} \}}{\sum_{j \in R(t_i)} \exp \{ \mathbf{x}'_j \hat{\boldsymbol{\beta}} \}}, \quad (9)$$

em que  $R(t_i)$  é o conjunto dos indivíduos em risco.

A forma padronizada dos resíduos de Schoenfeld frequentemente usada (Colosimo & Giolo, 2006), é definida por:

$$s_i^* = [\mathcal{I}(\hat{\boldsymbol{\beta}})]^{-1} \times \mathbf{r}_i \quad (10)$$

em que  $\mathcal{I}(\hat{\boldsymbol{\beta}})$  é a matriz de informação observada. A partir destes resíduos é possível obter o coeficiente de correlação de Pearson dos resíduos com cada covariável e também avaliar, por meio de testes de hipótese, a suposição de riscos proporcionais do modelo de Cox.

Uma vez violada a suposição de riscos proporcionais, uma solução, caso a não proporcionalidade seja importante, é estratificar os dados de modo que a suposição de riscos proporcionais seja válida em cada estrato. A análise estratificada consiste em dividir os dados em  $m$  estratos, conforme a indicação de violação da

suposição. O modelo de riscos proporcionais é então expresso como:

$$\lambda(t|\mathbf{x}_{ij}) = \lambda_{0j}(t) \exp\{\mathbf{x}'_{ij}\boldsymbol{\beta}\}, \quad (11)$$

para  $j = 1, \dots, m$  e  $i = 1, \dots, n_j$  em que  $n_j$  é o número de observações no  $j$ -ésimo estrato. Note que  $\boldsymbol{\beta}$  é assumido ser comum para todos os estratos. Segundo Colosimo & Giolo (2006), a estratificação não cria nenhuma complicação na estimação do parâmetro  $\boldsymbol{\beta}$ , e as propriedades assintóticas dos estimadores são obtidas a partir dos estimadores do modelo não estratificado.

## 2.2 Modelo Linear Generalizado

O modelo Linear Generalizado (GLM) surgiu a partir do artigo de Nelder & Wedderburn (1972). Este trabalho permitiu abrir uma leque de opções para a distribuição da variável resposta do modelo regressão linear, fazendo com que ela pertença a família exponencial de distribuições. Além disso, o modelo GLM permitiu obter uma maior flexibilidade para a relação funcional entre a média da variável resposta e o preditor linear.

Para entender como caracterizar uma variável pertencente à família exponencial e a flexibilidade mencionada do modelo, considere  $Y_1, \dots, Y_n$  variáveis aleatórias independentes, então, dizemos que uma variável aleatória pertence a família exponencial se sua função de densidade ou probabilidade, puder ser escrita da seguinte forma:

$$f(y_i; \theta_i; \phi) = \exp\left\{\frac{1}{a(\phi)} [y\theta_i - b(\theta_i)] + c(y_i; \phi)\right\}, \quad (12)$$

em que  $a$ ,  $b$  e  $c$  são funções que assumem valores em subconjuntos dos reais, o parâmetro  $\phi$  é conhecido como parâmetro de escala e  $\theta_i$  é denominado como parâmetro canônico da distribuição. A média e a variância são expressas por:

$$E(Y_i) = \mu_i = \frac{db(\theta_i)}{d\theta_i} \quad (13)$$

e

$$Var(Y_i) = \frac{d^2 b(\theta_i)}{d\theta_i^2} \cdot a(\phi). \quad (14)$$

A estrutura básica para um modelo Linear Generalizado será representada por:

$$g(\mu_i) = \eta_i = \mathbf{X}_i \boldsymbol{\beta}$$

e

$$Y_i \sim f_{\theta_i}(y_i),$$

em que  $g$  é denominado de função de ligação da média com o preditor linear,  $f_{\theta_i}(y_i)$  indica uma distribuição da família exponencial com parâmetro canônico  $\theta_i$ , sendo este parâmetro obtido através da equação (13).

A estimação dos parâmetros deste modelo é feita pelo método da máxima verossimilhança. Detalhes do processo de maximização da função de verossimilhança e aspectos computacionais podem ser visto Wood (2006).

### 2.3 Modelo Aditivo Generalizado

O modelo Linear Generalizado Aditivo (GAM) foi proposto por Hastie & Tibshirani (1986), sendo uma generalização do modelo Linear Generalizado (MLG). Em geral, no MLG estamos interessados em modelar a resposta média ou algum outro parâmetro de interesse como função linear  $\sum_{j=1}^p \beta_j X_j$  de um conjunto de covariáveis  $X_1, \dots, X_p$ . Esta função linear, que também chamada de preditor linear, nos modelos GAM's possui também uma função de suavização das covariáveis. O modelo tem a seguinte estrutura:

$$\eta(\mu_i) = \mathbf{X}_i^* \boldsymbol{\beta} + f_j(\mathbf{X}_i^*) = \mathbf{X}_i^* \boldsymbol{\beta} + f_1(x_{1i}) + f_2(x_{2i}) + f_3(x_{3i}) + \dots, \quad (15)$$

em que  $\mu_i = E(Y_i)$  e  $Y_i$ , a variável resposta, possui alguma distribuição da família exponencial, e  $\mathbf{X}_i^*$  é uma linha da matriz de covariáveis,  $\boldsymbol{\beta}$  é o vetor de parâmetros associados às covariáveis, e  $f_j$  são funções de suavização das covariáveis.

As funções  $f_j$  permitem flexibilizar a dependência da variável resposta com as covariáveis. De acordo com Wood (2006), esta flexibilidade traz dois problemas de como representar a função de suavização e quão suave ela deve ser. Para

obter a função  $f_j$  há diversas formas, mas todas tem como característica comum que a suavização em um ponto depende de suas observações e das observações vizinhas. As técnicas existentes são: regressão polinomial local com pesos (lowess), kernel, splines e splines com penalização (psplines). Detalhes de todas essas técnicas podem ser encontrados em Hastie & Tibshirani (1990) e Wood (2006).

Os splines são funções polinomiais descontínuas que impõem restrições sobre os pontos de uniões, denominados de nós, os quais dividem o intervalo dos valores da covariável de interesse, que será feita a suavização. Os splines dependem basicamente de três elementos:

- Grau do polinômio;
- Número de nós;
- Localização dos nós.

Para entender melhor como funciona os splines, considere uma covariável  $x \in [a, b]$ , e dois nós  $x_1$  e  $x_2$  que pertencem ao intervalo  $[a, b]$ . Se os nós definem um spline cúbico, temos as seguintes funções de base  $w_1(x) = 1$ ,  $w_2(x) = x$ ,  $w_3(x) = x^2$  e  $w_4(x) = x^3$ . Assim, um spline cúbico ajustará um polinômio de grau 3 em cada intervalo  $[a, x_1]$ ,  $(x_1, x_2]$  e  $(x_2, b]$ . Como não há garantia de continuidade da função de suavização nos nós, é necessário criar mais duas outras funções para compor a base:  $w_5(x) = (x - x_1)_+^3$  e  $w_6(x) = (x - x_2)_+^3$ , com  $k = 1, 2$ , tem-se:

$$(x - x_k)_+^3 = \begin{cases} (x - x_k)^3, & \text{se } x > x_k \\ 0, & \text{c.c.} \end{cases}$$

A partir dessas considerações, o spline cúbico pode escrito como uma combinação linear das funções de base:

$$f(x) = \sum_{j=1}^6 \beta_j w_j(x),$$

sendo os polinômios descrito da seguinte forma:

$$f(x) = \begin{cases} \beta_1 + \beta_2 x + \beta_3 x^2 + \beta_4 x^3, & \text{se } a \leq x \leq x_1 \\ \beta_1 + \beta_2 x + \beta_3 x^2 + \beta_4 x^3 + \beta_5 (x - x_1)^3, & \text{se } x_1 < x \leq x_2 \\ \beta_1 + \beta_2 x + \beta_3 x^2 + \beta_4 x^3 + \beta_5 (x - x_1)^3 + \beta_6 (x - x_2)^3, & \text{se } x_2 < x \leq b \end{cases}$$

A solução do problema de estimação dos coeficientes  $\beta_j$  das funções de splines, se concentra em minimizar a soma de quadrados penalizados:

$$\sum_{i=1}^n [y_i - f(x_i)]^2 + \lambda \int \left( \frac{d^2 f(x)}{dx^2} \right)^2 dx, \quad (16)$$

em que  $\lambda$  é o parâmetro de penalização que controla o ajuste e a suavidade de  $f$ . A primeira parcela do somatório serve para aproximar a função de suavização dos dados e a segunda parcela penaliza a curvatura.

No caso em que a função de suavização possui duas covariáveis, como é o caso de localizações espaciais que dependem da latitude e longitude, a função objetiva na equação (16), pode ser escrita como:

$$\sum_{i=1}^n [y_i - f(x_{1i}, x_{2i})]^2 + \lambda \left\{ \int \int \left[ \left( \frac{\partial^2 f(x)}{\partial x_1^2} \right)^2 + 2 \left( \frac{\partial^2 f(x)}{\partial x_1 \partial x_2} \right)^2 + \left( \frac{\partial^2 f(x)}{\partial x_2^2} \right)^2 \right] dx_1 dx_2 \right\}. \quad (17)$$

Uma extensão do modelo de Cox, apresentado na equação (1), utilizando as funções de suavização é apresentada por Hastie & Tibshirani (1990). A ideia é reescrever a equação (1), da seguinte forma:

$$\lambda(t|\mathbf{x}_i) = \lambda_0(t) \exp \{ \eta(\mathbf{X}) \}, \quad (18)$$

em que  $\eta(\mathbf{X}) = \sum_{j=1}^p f_j(X_j)$ , como apresentado na equação (15).

Sejam  $\eta_i = \eta(x_i)$  e  $\boldsymbol{\eta} = (\eta_1, \dots, \eta_n)'$ , com  $i = 1, \dots, n$  indivíduos,  $t_r$  o  $r$ -ésimo tempo de falha, com  $r = 1, \dots, k$ , a verossimilhança parcial da equação (18) será expressa por:

$$L(f_1, \dots, f_p) = \prod_{r=1}^k \frac{\exp(\sum_{i \in D_r} \eta_i)}{[\sum_{i \in R_r} \exp(\eta_i)]^{d_r}}, \quad (19)$$

em que  $R_r$  é o conjunto dos indivíduos que estão sob risco até o tempo  $t_r$ ,  $D_r$  é o conjunto de falhas no tempo  $t_r$  e  $d_r$  é o número de falhas no tempo  $t_r$ .

Aplicando o logaritmo na equação (19), obtêm-se:

$$l(f_1, \dots, f_p) = \sum_{r=1}^k \left\{ \sum_{i \in D_r} \eta_i - d_r \cdot \log \left[ \sum_{i \in R_r} \exp(\eta_i) \right] \right\}. \quad (20)$$

Observe que, na equação (20), se considerarmos a probabilidade de falha do indivíduo  $i$  no tempo  $t_r$ , sem a ocorrência empates, teríamos:

$$l(f_1, \dots, f_p) = \eta_i + \log \left[ \sum_{i \in R_r} \exp(\eta_i) \right],$$

que equivale à família exponencial com parâmetro  $b(f_j) = \log [\sum_{i \in R_r} \exp(\eta_i)]$ . A média condicional e a função de variância como nos GLM, podem ser obtidos por  $b'(f_j)$  e  $b''(f_j)$ .

Para estimação dos parâmetros do modelo é necessário obter as derivadas parciais  $\frac{\partial l}{\partial \eta_i}$ ,  $\frac{\partial^2 l}{\partial^2 \eta_i}$  e  $\frac{\partial^2 l}{\partial \eta_i \partial \eta_{i'}}$ , que serão utilizadas no problema de minimizar a expressão apresentada na equação (16). Detalhes computacionais podem ser visto em Hastie & Tibshirani (1990).

## 2.4 Inferência Bayesiana

Os métodos bayesianos tem crescido nos últimos vinte anos devido aos resultados de convergência e métodos computacionais eficientes com objetivo de obter distribuições posteriores dos dados (Carlin & Louis, 2009). Na análise bayesiana, são feitas inferências dos dados usando modelos de probabilidade para quantidades observadas e para quantidades que desejamos conhecer, sendo esta última, denotada pelo parâmetro  $\theta$ . A informação sobre a quantidade observada está presente na amostra aleatória dos dados, denotada por  $Y$ . Além disso, considera-se um conhecimento a priori de  $\theta$  que é resumido pela distribuição a priori do parâmetro  $p(\theta)$ . O peso da informação da amostra ou da informação a priori é ajustado em detrimento de quão informativa é a distribuição a priori atribuída ao parâmetro. Através do teorema de

Bayes, podemos combinar ambas as informações, resultando na seguinte distribuição a posteriori:

$$p(\theta|y) = \frac{p(y|\theta)p(\theta)}{p(y)} \implies p(\theta|y) \propto p(y|\theta)p(\theta) \quad (21)$$

Estas informações contidas na equação (21), podem ser resumidas da seguinte forma:

1.  $p(\theta)$  é a distribuição a priori para o vetor de parâmetros  $\theta$ , usando portanto a distribuição de probabilidade sobre os parâmetros como uma forma de refletir a informação (incerteza/ignorância) sobre  $\theta$  antes de coletar os dados;
2.  $p(y|\theta) = L(y|\theta)$  é a verossimilhança, que contém toda a informação sobre  $\theta$  proveniente dos dados  $y$  a ser utilizada;
3.  $p(y) = L(y)$  é a verossimilhança marginal de  $y$ , que em métodos como MCMC (Markov Chain Monte Carlo), pode ser omitida na obtenção da posteriori.
4.  $p(\theta|y)$  é a distribuição conjunta a posteriori que expressa a informação (incerteza) sobre o vetor de parâmetros após combinar a informação a priori e aquela contida na verossimilhança.

Na estatística, inicialmente, sempre buscamos resumir numericamente a informação dos dados através de uma medida de tendência central (média, mediana, moda). No cenário bayesiano, a média pode ser obtida, considerando, por exemplo, que o parâmetro de interesse  $\theta$  assuma valores em um espaço paramétrico contínuo  $\Theta$ , de modo que a média posteriori é igual a:

$$E(\theta|y) = \int_{\theta \in \Theta} \theta \cdot p(\theta|y) d\theta. \quad (22)$$

Como na abordagem frequentista ao utilizar estimadores pontuais do parâmetro de interesse, sabemos que eles não fornecem uma medida de precisão, isto é, não demonstram o quanto poderíamos estar errado ao utilizar esse único valor



para estimar o parâmetro. Uma forma de considerar esta precisão e apresentar uma estimativa mais informativa, na abordagem frequentista, é através do intervalo de confiança. A interpretação, na abordagem frequentista, sugere que no intervalo de confiança  $100 \cdot (1 - \alpha)\%$ , se repetirmos o mesmo experimento sob as mesmas condições, o verdadeiro valor parâmetro  $\theta$ , não estará contido em apenas  $\alpha \cdot 100\%$  das vezes. Na abordagem bayesiana devido à obtenção da distribuição posteriori, temos o intervalo de credibilidade (C), em que a interpretação é feita em termos de probabilidade, ou seja, tem-se uma probabilidade posteriori de  $100 \cdot (1 - \alpha)\%$  do verdadeiro valor de  $\theta$  estar dentro dos limites de C,  $p(\theta \in C|y)$ .

No âmbito da pesquisa, há sempre o interesse em explicar e entender a variável resposta do estudo em função de variáveis que expliquem ou influenciem mudanças na variável resposta. Para esta finalidade, recorreremos à modelagem estatística que, dependendo do problema tratado, o processo de modelagem pode se tornar bastante complexo.

Em epidemiologia, vários dados são caracterizados por uma estrutura espaço-temporal e, geralmente, a variável resposta pode não apresentar distribuição normal, o que na prática pode inviabilizar o uso de modelos de regressão linear. Neste contexto, os modelos gaussianos latentes são modelos flexíveis para várias aplicações, que abrangem os modelos lineares generalizados, modelos lineares generalizados aditivos, modelos de regressão semi-paramétricos e modelos espaço-temporal (Rue et al., 2009). A classe de modelos gaussianos latentes servem para modelar efeitos não lineares das covariáveis, heterogeneidade de grupos específicos e a dependência entre o espaço e o tempo. Em geral, para esses modelos a distribuição posteriori não é avaliada analiticamente, fazendo uso do método MCMC. Porém em virtude da quantidade de processo inferenciais e do tempo computacional para realizar as simulações até a convergência, o MCMC pode ser inviável na prática (Rue et al., 2009). O INLA - Integrated Nested Laplace Aproximation, oferece uma alternativa para contornar limitações do MCMC, retomando as ideias de integração numérica, porém agora, em um cenário Bayesiano, onde aproximações determinística são usadas para obter

distribuições posteriores (Ribeiro Jr et al., 2012).

Um maneira geral de especificar os modelos Gaussianos latentes, utilizados no INLA, é modelar a média para a  $i$ -ésima unidade, através de um preditor linear aditivo.

$$\eta_i = \beta_0 + \sum_{j=1}^{p-1} \beta_j x_{ji} + \sum_{j=1}^L f_j(x_{ji}), \quad (23)$$

$\beta_0$ : intercepto aleatório;

$\boldsymbol{\beta}$ : vetor de coeficientes;

$\mathbf{X}$ : vetor de covariáveis;

$\mathbf{f} = (f_1(\cdot), \dots, f_L(\cdot))$  é uma coleção de funções de covariáveis de  $\mathbf{X}$ , que podem acomodar efeitos espaciais.

## 2.5 Método de estimação INLA

O INLA é um algoritmo proposto por Rue et al. (2009), utilizado especificamente para modelos gaussianos latentes. O objetivo do método é realizar aproximações determinísticas que são usadas para obter distribuições posteriores dos parâmetros desses modelos. O tipo de aproximação determinísticas são as aproximações de Laplace. Blangiardo & Cameletti (2015), apresenta uma maneira simples de entender as aproximações de Laplace. Considere que estamos interessados em calcular a seguinte integral:

$$\int f(y) dy = \int \exp \{ \log f(y) \} dy, \quad (24)$$

em que  $f(y)$  é uma função de densidade de uma variável aleatória  $Y$ . Representando a função  $\log f(y)$  pela expansão em série de Taylor em torno de  $y = y_0$ , obtemos:

$$\log f(y) \approx \log f(y_0) + (y - y_0) \frac{d \log f(y)}{dy} \Big|_{y=y_0} + \frac{(y - y_0)^2}{2} \frac{d^2 \log f(y)}{dy^2} \Big|_{y=y_0}. \quad (25)$$

Se  $y_0^*$  é definido de modo que seja igual ao máximo da função  $\log f(y)$ , temos que  $\frac{d \log f(y)}{dy} \Big|_{y=y_0^*} = 0$ , e assim a equação (25), será igual a:

$$\log f(y) \approx \log f(y_0^*) + \frac{(y - y_0^*)^2}{2} \frac{d^2 \log f(y)}{dy^2} \Big|_{y=y_0^*}, \quad (26)$$

substituindo a expressão (26) na integral da equação (24), obtêm-se:

$$\int f(y) dy \approx \int \exp \left\{ \log f(y_0^*) + \frac{(y - y_0^*)^2}{2} \frac{d^2 \log f(y)}{dy^2} \Big|_{y=y_0^*} \right\} dy = \int f(y_0^*) \exp \left\{ \frac{(y - y_0^*)^2}{2} \frac{d^2 \log f(y)}{dy^2} \Big|_{y=y_0^*} \right\} dy,$$

associando a integral com a função de densidade da distribuição normal, temos que  $\frac{d^2 \log f(y)}{dy^2} = -\frac{1}{\sigma^2}$ , e assim a integral pode ser reescrita como:

$$\int f(y) dy \approx \int f(y_0^*) \exp \left\{ -\frac{(y - y_0^*)^2}{2} \frac{1}{\sigma^{2*}} \right\} dy.$$

No caso da integral ser avaliada em um intervalo  $(a, b)$ , a integral definida é aproximada por:

$$\int_a^b f(y) dy \approx \int_a^b f(y_0^*) \exp \left\{ -\frac{(y - y_0^*)^2}{2} \frac{1}{\sigma^{2*}} \right\} dy = f(y_0^*) \sqrt{2\pi\sigma^{2*}} (\Phi(b) - \Phi(a)).$$

Para o vetor de respostas  $\mathbf{y} = (y_1, \dots, y_n)$  com preditor linear aditivo apresentado na equação (23), e definido o espaço paramétrico  $\Theta = \{\alpha, \beta, \mathbf{f}\}$  e hiperparâmetros  $\psi = \{\psi_1, \dots, \psi_k\}$ , o objetivo principal do INLA, como pode ser visto em mais detalhes em Blangiardo & Cameletti (2015) e Rue et al. (2009), é aproximar a marginal posteriori por:

$$\tilde{p}(\theta_i|\mathbf{y}) \approx \int \tilde{p}(\theta_i|\psi, \mathbf{y}) d\psi,$$

em que  $\tilde{p}()$  é a aproximação gaussiana dada pelo método INLA.

## 2.6 Modelo espaço-temporal na análise de dados de Sobrevida

Padrões de mortalidade ou morbidade, propagação de epidemias, transmissão sexual de doenças ou transferências de comportamentos ou valores não podem ser explicados sem uma abordagem que, além de considerar os grupos, estude o espaço e o tempo (Carvalho & Souza-Santos, 2005). A mesma literatura aponta que a maioria das pesquisas em epidemiologia estão pautadas no risco relativo do

indivíduo, ao passo que a estatística espacial pode fornecer subsídios de aglomerados espaciais ou espaço-temporais, na avaliação e monitoramento ambiental e aplicados ao planejamento e avaliação de uso de serviços de saúde.

No contexto de análise de sobrevivência, além dos aspectos levantados anteriormente, é necessário uma modelagem de estatística espacial que leve em conta o tempo até a ocorrência do evento de interesse, as variáveis explicativas e a informação espacial dos casos. O modelo semi-paramétrico bayesiano que foi implementado por Martino & Akerkar (2010) no programa R (R Development Core Team, 2016), pacote INLA, é um modelo que contempla todas essas características.

Para este modelo, os tempos até a ocorrência do evento de interesse são particionados em um intervalo finito,  $0 = s_0 < s_1 < \dots < s_K$ ,  $t_i < s_k$ , para  $\forall t_i$ , com  $i = 1, \dots, n$ , e que o risco basal seja constante no intervalo de tempo  $(s_{k-1}, s_k]$ , com  $k = 1, \dots, K$ , ou seja,  $\lambda_0(t) = \lambda_k$ , para cada  $t \in (s_{k-1}, s_k]$ . O modelo para o  $i$ -ésimo indivíduo no  $k$ -ésimo intervalo de tempo, pertencente ao  $j$ -ésimo setor censitário é dado por:

$$\lambda_{ikj}(t) = \lambda_0(t) \exp \{ \mathbf{x}'_i \boldsymbol{\beta} + f(z_j) \} = \exp \{ \mathbf{x}'_i \boldsymbol{\beta} + f(z_j) + b_k \} = \exp \{ \boldsymbol{\eta}_{ik} \}, \quad (27)$$

em que  $\lambda_{ikj}(t)$  representa a função taxa de risco,  $b_k = \log [\lambda_0(t)]$ ,  $\boldsymbol{\beta} = (\beta_1, \dots, \beta_p)'$  é o vetor de dimensão  $1 \times p$  de parâmetros associados às covariáveis,  $\mathbf{x}_i$  é o vetor de dimensão  $p \times 1$  de covariáveis observadas para o  $i$ -ésimo indivíduo e  $f(z_j)$  é uma função de suavização do efeito espacial no  $j$ -ésimo setor censitário. Para o componente  $f(z_j)$ , adotou-se um modelo Intrínseco Condicional Autoregressivo – ICAR (Besag & Kooperberg, 1995). Este modelo consiste de um caso particular dos modelos Condicionais Autoregressivos – CAR, que são usados para dados agregados por unidade de área, cuja a relação entre as unidades ocorre em termos da adjacência entre elas. O modelo ICAR é expresso da seguinte forma:  $f(z_j) | f(z_w) \sim N \left( \frac{1}{l_j} \sum_{j \sim w} f(z_w), \frac{1}{l_j \tau_z} \right)$ , em que  $l_j$  é o número de setores que compartilham a mesma fronteira com o setor  $j$ ,  $j \sim w$  indica que dois setores censitários  $j$  e  $w$  são vizinhos e  $\tau_z$  é o parâmetro de variância do modelo.

No cenário bayesiano, para realizar as inferências do modelo da equação (27), é possível a utilização do método INLA.

## 2.7 AIDS

O HIV é a sigla em inglês do vírus da imunodeficiência humana. Este vírus possui como material genético RNA e uma enzima denominada transcriptase reversa, capaz de converter RNA do vírus em DNA. As células mais prejudicadas por esse vírus são os linfócitos T CD4+, células de defesa do organismo. Alterando o DNA dessas células, o HIV se multiplica, e em seguida rompe os linfócitos em busca de outras células T CD4+, para continuar a infecção. Como o HIV é mutagênico, o sistema imunológico tem dificuldade de organizar uma resposta imune a esse vírus, até o momento que as células de defesa começam a funcionar com menor eficiência e serem destruídas. Neste momento ocorre uma redução drástica dos linfócitos T CD4+, chegando o indivíduo a ficar com menos de 200 cel./mm<sup>3</sup> de sangue. Este estágio, ocasionado pelo HIV, é denominado como síndrome da imunodeficiência adquirida (AIDS), em que o organismo fica mais vulnerável a diversas doenças oportunistas como tuberculose e até câncer. As formas de transmissão do HIV mais comuns ocorrem por meio de relações sexuais, transfusão sanguínea, transmissão vertical (da mãe para o recém-nascido) e seringas contaminadas. Quanto antes o indivíduo for diagnosticado e iniciar o tratamento, maior será a sua expectativa de vida (Ministério da Saúde, 2017).

Atualmente 35 milhões de pessoas vivem com o HIV no mundo e 19 milhões não sabem que são HIV-positivos. Um em cada quatro novos casos é de mulheres jovens da África Subsaariana e a população mais vulnerável para o HIV são de prisioneiros, e também para outras infecções como tuberculose, hepatite B e C. No mundo 22 milhões pessoas não têm acesso ao tratamento para a AIDS, sendo que é possível o fim da epidemia da AIDS desde que diminua as diferenças de acesso ao tratamento entre os países ricos e pobres. (UNAIDS, 2014)

Estima-se que o número de pessoas com 50 anos ou mais que vivem

com o HIV/AIDS é igual a 4,2 milhões, e todos os anos, quase 120.000 pessoas nesta faixa etária adquirem este vírus (UNAIDS, 2014). Como o aumento na proporção de idosos vem crescendo nas últimas décadas é importante que eles tenham cuidados especializados para o HIV e outras doenças crônicas.

O HIV em países industrializados foi vinculado, inicialmente, a homens que fazem sexo com homens e disseminou-se rapidamente, sendo que ao longo das últimas décadas vem afetando vários segmentos da população, principalmente adultos jovens, com um grande impacto na economia, na sociedade e na estrutura familiar (Chequer et al., 1992; Guibu et al., 2011).

Apesar dos números de infecção do vírus do HIV causar preocupação, vários estudos apontam uma redução drástica da mortalidade por AIDS no Brasil e em várias partes do mundo, em decorrência da utilização da terapia antirretroviral altamente ativa (HAART). No Brasil, o acesso à medicação antirretroviral tem sido garantido pelo Ministério da Saúde desde o ano de 1996, o que proporcionou que o tempo de sobrevivência dos pacientes aumentasse sensivelmente após o início da disponibilização da HAART. Além disso, observa-se uma redução no número de internações de pessoas com AIDS, diminuição de infecções oportunistas e do aumento de doenças crônicas (Chequer et al., 1992; Guibu et al., 2011; Lima et al., 2007; Menesia et al., 2001).

De acordo com o Boletim Epidemiológico HIV/Aids (2016), desde o início da epidemia de AIDS no Brasil, até junho de 2016 foram registrados 842.710 casos de AIDS. Sendo que nos primeiros anos a epidemia se concentrava mais nas capitais da região sudeste e em alguns municípios do estado de São Paulo. No Brasil as regiões sudeste, sul, nordeste, centro-oeste e norte, correspondem, respectivamente, a 53,0%, 20,1%, 15,1%, 6,0% e 5,9% do total de casos identificados de 1980 até junho de 2016. Apesar de a região sudeste apresentar a maior porcentagem dos casos, o boletim aponta que esta região é a única a apresentar tendência de queda na detecção de novos casos nos últimos 10 anos.

No Brasil a maior concentração de casos de AIDS ocorre na proporção

de 65,1% para homens e 34,9% para mulheres. Em relação à faixa etária, a maior concentração está na faixa de 25 a 39 anos para ambos os sexos, sendo que entre os homens tem-se 53,0% e entre as mulheres 49,4% do total de casos registrados de 1980 a junho de 2016. Além disso, ocorre predomínio da categoria de exposição heterossexual e a proporção de usuários de drogas vem diminuindo ao longo dos anos em todo o Brasil. Na atualidade, observa-se que a epidemia atinge vários grupos da sociedade sofrendo um processo de heterossexualização, feminização e pauperização (Santos et al., 2002). Em relação ao óbito foram registrados, desde 1980 até dezembro de 2014, 303.353 óbitos, sendo 60,3% na região sudeste. Em contrapartida, a maior redução nos últimos anos ocorreu na região sudeste, com uma redução de 20,7% dos óbitos.

No estado de São Paulo, com base no Boletim Epidemiológico CRT-PE – DST/AIDS/CVE (2015), foram registrados 244.641 casos de AIDS. A análise epidemiológica descrita no boletim aponta uma queda na taxa de incidência de 30% nos casos de AIDS nos últimos 10 anos e uma redução de 72% nos últimos 20 anos, embora de forma heterogênea, ainda permaneça entre as cinco principais causas de mortes no estado de São Paulo. O panorama da epidemia paulista desenha-se como evidência preocupante entre jovens, homens que fazem sexo com homens, pessoas que usam drogas e profissionais do sexo. Em relação à mortalidade no período 1985 a 2014 foram registrados 108.820 óbitos no estado de São Paulo. As taxas vêm diminuindo ao longo dos anos, apresentando como maior taxa 22,9 mortes por 100.000 habitantes em 1995 e chegando a 6,4 mortes por 100.000 habitantes em 2014.

## 3 MATERIAL E MÉTODOS

### 3.1 Dados

O estudo envolveu três coortes retrospectivas de pacientes notificados por AIDS, moradores do município de Campinas, a nordeste do estado de São Paulo (aproximadamente 1 milhão de habitantes). Os dados foram obtidos do Sistema de Informação de Agravos de Notificação (SINAN) e dos óbitos que constam no banco de dados do Sistema de Informação sobre Mortalidade (SIM) de moradores de Campinas. Uma coorte foi composta por indivíduos notificados nos primeiros anos da epidemia, de 1980 a 1990 (coorte 1), outra com as notificações de 1996 a 2000 (coorte 2) e uma terceira no período de 2001 a 2005 (coorte 3). Nestas coortes, os tempos de seguimento dos indivíduos foram ao redor de 10 anos, respectivamente, até os anos 2000, 2010 e 2014, afim de identificar a ocorrência do óbito. Não foram incluídos pacientes que não constavam na lista de notificação do SINAN. Após checagem de duplicações e inconsistência dos bancos, as variáveis utilizadas no estudo foram: sexo, idade com que o indivíduo foi diagnosticado com o vírus do HIV, escolaridade, categoria de transmissão (homossexual/bissexual ou heterossexual, uso de droga injetáveis), contagem de células linfócitos T CD4+, tuberculose (todas as formas clínicas), câncer, (pelo menos um dos seguintes diagnósticos: linfoma não Hodgkin, linfoma primário do cérebro, câncer uterino ou Sarcoma de Kaposi), além disso também foi obtida a localização das residências georreferenciadas nos setores censitários. Para agregar as localizações das residências dos indivíduos que pertencem ao mesmo setor censitário, utilizamos o software QGIS 2.8.2 e o mapa da cidade de Campinas-SP, obtido no site do IBGE, utilizando a Malha de Setor Censitário



Urbano do ano 2000. Todas as análises estatísticas foram realizados no programa R (R Development Core Team, 2016).

Na coorte 1 havia um total de 286 indivíduos, sendo 245 (85,66%) homens e 41 (14,34%) mulheres. Na coorte 2, os dados correspondiam a 1456 indivíduos, consistindo em 450 (30,91%) mulheres e 1006 (69,09%) homens. Na coorte 3, observou-se 1342 indivíduos, sendo 831 (61,92%) homens e 511 (38,08%) mulheres.

## Artigo 1

Aceito para a publicação na Revista Brasileira de Epidemiologia – RBE

### RISCO ESPACIAL DE ÓBITO DE PACIENTES COM AIDS EM CAMPINAS, SÃO PAULO, BRASIL

SPATIAL RISK OF DEATHS IN PATIENTS WITH AIDS IN CAMPINAS, SÃO  
PAULO, BRAZIL

Thiago Santos Mota<sup>1</sup>, Maria Rita Donalizio<sup>2</sup>, Liciano Vaz de Arruda Silveira<sup>3</sup>

Departamento de Bioestatística. Instituto de Biociências – UNESP/Botucatu<sup>1,3</sup>

Departamento de Saúde Coletiva. Faculdade de Ciências Médicas – UNICAMP/Campinas<sup>2</sup>

## RESUMO

**Objetivo:** O estudo teve como objetivo analisar o risco espacial de mortalidade por AIDS em setores censitários de município brasileiro de grande porte. **Método:** Foram estudadas três coortes retrospectivas de indivíduos notificados por AIDS e residentes no município de Campinas, São Paulo, nos períodos de 1980 a 1990 (coorte 1), 1996 a 2000 (coorte 2) e outra com indivíduos notificados de 2001 a 2005 (coorte 3), afim de identificar a ocorrência do óbito. Foi ajustado um modelo semi-paramétrico bayesiano, empregando o método de Aproximação de Laplace Aninhada e Integrada (INLA), que permitiu obter mapas de risco de óbito nas três coortes estudadas. **Resultados:** A incidência de óbito nas coortes 1, 2 e 3 foram, respectivamente, 72,73%, 32,21% e 13,11%. Os mapas de risco de óbito apontaram tendência de queda da mortalidade e evidenciaram os setores que apresentaram maior e menor risco em cada período. O modelo mostrou que nas três coortes, os fatores associados ao maior risco de óbito foram: ser do sexo masculino, possuir idade de diagnóstico superior a 49 anos. A categoria de transmissão homossexual/bissexual foi associada ao menor risco de morrer. **Conclusões:** Mesmo diante da diminuição do risco de morte por AIDS nos últimos anos, após o acesso à terapia antirretroviral altamente ativa, a maior incidência se concentra entre os mais pobres nas três coortes estudadas. Os mapas e o fatores de risco obtidos orientam possíveis ações e monitoramento da doença no município.

**Palavras-chave:** AIDS. Mortalidade. Terapia antirretroviral. Análise de Sobrevida

## ABSTRACT

**Objective:** The objective of this study was to analyze the spatial risk of AIDS mortality in census tracts in a large Brazilian municipality. **Method:** We studied three retrospective cohorts of individuals with AIDS and residents in the city of Campinas, state of São Paulo. The occurrence of death was analyzed from 1980 to 1990 (cohort 1), 1996 to 2000 (cohort 2) and another with individuals reported from 2001 to 2005 (cohort 3). It was adjusted a Bayesian semi-parametric model, using the Integrated Nested Laplace Approximation - INLA method, which allowed the mapping of mortality risk for the three cohorts. **Results:** The incidence of death in cohorts 1, 2 and 3 were, respectively, 72.73%, 32.21% and 13.11%. The risk maps showed a tendency of mortality decrease and presented the sectors with higher and lower risk in each period. The model showed that in the three cohorts, the factors associated with the highest risk of death were: male, with a diagnosis age greater than 49 years. The homosexual/bisexual category of transmission was associated with the lower risk of dying. **Conclusions:** Even in the face of a reduction in the risk of death from AIDS in recent years, after access to highly active antiretroviral therapy, the highest incidence is concentrated among the poorest in the three cohorts studied. The maps and the risk factors obtained guide possible actions for monitoring the disease in the municipality. Integrated Nested Laplace Approximation.

**Key words:** AIDS. Mortality. HAART. Survival Analysis

## INTRODUÇÃO

A epidemia da AIDS ao longo das últimas três décadas vem afetando vários segmentos da população, principalmente adultos jovens, com um grande impacto na economia, na qualidade de vida dos indivíduos.<sup>1</sup> Vários estudos apontam uma redução drástica da mortalidade por AIDS em várias partes do mundo, em decorrência da utilização da terapia anti-retroviral altamente ativa (HAART).<sup>1,2,3,4</sup> No Brasil, o acesso à medicação antirretroviral combinada tem sido garantido pelo Ministério da Saúde desde o ano de 1996. O seu impacto no aumento da sobrevivência dos pacientes, na redução de internações, de infecções oportunistas e no aumento de doenças crônicas tem sido registrado em vários estudos no país.<sup>3,5,6,7</sup> Abordagens geoespaciais têm sido utilizadas para estudar tendências da epidemia de AIDS no Brasil, apontando disparidades regionais nos riscos de adoecer e morrer.<sup>8,9</sup> No estado de São Paulo, na década de 1980, a doença acometia principalmente homossexuais, usuários de drogas injetáveis e pacientes transfundidos. Nas décadas seguintes observou-se que a epidemia atingiu vários grupos da sociedade, sofrendo um processo de heterossexualização, feminização e pauperização. Atualmente a doença tem se expandido em adultos jovens com baixo nível socioeconômico.<sup>10,11</sup> Nesta perspectiva, com intuito de entender a dinâmica da epidemia de AIDS no município de Campinas-SP, Brasil, foram georreferenciados os casos notificados pelo Sistema de Informação de Agravos de Notificação (SINAN) e dos óbitos que constam no banco de dados do Sistema de Informação sobre Mortalidade (SIM) de moradores da cidade, com o objetivo de analisar o risco espacial de óbito entre pacientes notificados por AIDS nos setores censitários deste município, ao longo do tempo.

## MÉTODOS

O estudo envolveu três coortes retrospectivas de pacientes maiores de 13 anos, notificados por AIDS pelo Sistema de Informação de Agravos de Notificação (SINAN), moradores do município de Campinas, a nordeste do estado de São Paulo (aproximadamente 100 km de distância de São Paulo).

madamente 1 milhão de habitantes). Foram excluídos 1091 (26,13%) pacientes cuja notificação coincidiu com a data do óbito para reduzir o impacto dos tempos destes pacientes na sobrevida, centrando os resultados dos estimadores nos tempos dos pacientes com acesso ao seguimento e tratamento.<sup>6</sup> Uma coorte foi composta por indivíduos notificados nos primeiros anos da epidemia para identificar a ocorrência de óbito, de 1980 a 1990 (coorte 1 – seguimento de 10 anos, com um tempo médio de sobrevida dos indivíduos de 5,7 anos, com um desvio padrão de 6,5 anos), outra com as notificações de 1996 a 2000 (coorte 2 – seguimento de 10 anos, com um tempo médio de sobrevida de 9,2 anos, com um desvio padrão de 5,6 anos) e uma terceira no período de 2001 a 2005 (coorte 3 – seguimento de 7 a 9 anos, com um tempo de sobrevida médio de 8,9 anos, com um desvio padrão de 3 anos). Além do SINAN, foram obtidas informações no Sistema de Informação de Mortalidade sobre os óbitos dos pacientes notificados. Não foram incluídos pacientes que não constavam na lista de notificação do SINAN. Após checagem de duplicações e inconsistência dos bancos, as variáveis utilizadas no estudo foram: sexo, idade com que o indivíduo foi diagnosticado com o vírus do HIV, escolaridade, categoria de transmissão (homossexual/bissexual ou heterossexual, por uso de droga injetáveis), contagem de células linfócitos T CD4+ categorizada em contagens menores ou iguais a 350 cel./mm<sup>3</sup> de sangue como critério de AIDS e analisada no modelo como variável contínua, além disso também foi obtida a localização das residências georreferenciadas nos setores censitários. Na coorte 1 havia um total de 286 indivíduos, sendo 245 (85,66%) homens e 41 (14,34%) mulheres. Na coorte 2, os dados correspondiam a 1456 indivíduos, consistindo em 1006 (69,09%) homens e 450 (30,91%) mulheres. Na coorte 3, observou-se 1342 indivíduos, sendo 831 (61,92%) homens e 511 (38,08%) mulheres. Para modelar o risco de óbito, foi necessária uma abordagem que leva em conta as variáveis de interesse e a informação espacial dos casos. Utilizou-se o modelo de regressão semi-paramétrico em um cenário bayesiano, que permite modelar o risco de óbito em relação aos setores censitários. O modelo semi-paramétrico bayesiano utilizado está implementando no programa R, pacote INLA.<sup>12,13</sup> A suposição

foi que os tempos até a ocorrência óbito sejam particionados em um intervalo finito,  $0 = s_0 < s_1 < \dots < s_K$ , para  $\forall t_i$ , com  $i = 1, \dots, n$ , e que o risco basal seja constante no intervalo de tempo  $(s_{k-1}, s_k]$ , com  $k = 1, \dots, K$ , ou seja,  $\lambda_0(t) = \lambda_k$ , para cada  $t \in (s_{k-1}, s_k]$ . O modelo para o  $i$ -ésimo indivíduo no  $k$ -ésimo intervalo de tempo, pertencente ao  $j$ -ésimo setor censitário é dado por:

$$\lambda_{ikj}(t) = \lambda_0(t) \exp \{ \mathbf{x}'_i \boldsymbol{\beta} + \mathbf{f}(\mathbf{z}_j) \} = \exp \{ \mathbf{x}'_i \boldsymbol{\beta} + \mathbf{f}(\mathbf{z}_j) + \mathbf{b}_k \} = \exp \{ \boldsymbol{\eta}_{ikj} \}$$

em que  $\lambda_{ikj}(t)$  representa a função taxa de risco,  $b_k = \log \{ \lambda_0(t) \}$ ,  $\boldsymbol{\beta} = (\beta_1, \dots, \beta_p)'$  é o vetor de dimensão  $1 \times p$  de parâmetros associados às covariáveis,  $\mathbf{x}'_i$  e o vetor de dimensão  $p \times 1$  de covariáveis observadas para o  $i$ -ésimo indivíduo e  $f(z_j)$  é a função do efeito espacial no  $j$ -ésimo setor censitário da cidade de Campinas-SP. Para o componente  $f(z_j)$ , adotou-se um modelo ICAR<sup>14</sup>, que pode ser expresso por:  $f(z_j) | f(z_w) \sim N \left( \frac{1}{l_j} \sum_{j \sim w} f(z_w), \frac{1}{l_j \tau_z} \right)$ . Em que  $l_j$  é o número de setores que compartilham a mesma fronteira com o setor  $j$ ,  $j \sim w$  indica que dois setores censitários  $j$  e  $w$  são vizinhos. Para o vetor de parâmetros são assumidas priors gaussianas,  $\boldsymbol{\beta} \sim N(0; 10^4 \mathbf{I})$ , para o vetor  $\mathbf{b} = (b_1, \dots, b_K)$ , priors gaussianas,  $b_{k+1} - b_k \sim N(0, \tau_b^{-1})$ , de um passeio aleatório de primeira ordem (RW1) e para os hiperparâmetros,  $\tau_b$  e  $\tau_z$ , priors independentes com distribuição gama,  $\Gamma(1; 0, 001)$ . Para os intervalos de tempo foi utilizado  $K = 20$ . No cenário bayesiano, para realizar as inferências desse modelo, a abordagem de estimação Bayesiana INLA (Aproximação de Laplace Aninhada e Integrada) é uma alternativa para o MCMC (Markov Chain Monte Carlo).<sup>15</sup> Basicamente, a metodologia INLA foi criada para uma classe de modelos de regressão aditivos estruturados. A ideia central da metodologia INLA é fazer integrações numéricas em um contexto Bayesiano, em que apenas aproximações determinísticas são usadas para obter as distribuições a posteriori, obtendo assim um ganho computacional expressivo.<sup>15</sup> Para agregar as localizações das residências dos indivíduos que pertencem ao mesmo setor censitário, utilizamos o software QGIS 2.8.2 e o mapa da cidade de Campinas-SP, obtido no site do IBGE, utilizando a Malha de Setor Censitário Urbano do ano 2000. Todas as análises es-

tatísticas foram realizados no programa R.<sup>13</sup> Para identificação de regiões da cidade de maior carência socioambiental, foi utilizado o mapa de distribuição espacial do índice de Carência Socioambiental (ICSA), segundo classes socioeconômicas e ambientais por área de cobertura dos Centros de Saúde em Campinas desenvolvido por Costa et al.<sup>16</sup> Este mapa possibilita a localização de três estratos de baixa, moderada e alta carência socioambiental a partir de análise estatística de componentes principais de variáveis demográficas e socioambientais do Censo 2000 da Fundação Instituto Brasileiro de Geografia e Estatística (FIBGE).

## RESULTADOS

Inicialmente, realizou-se a descrição das variáveis nas três coortes estudadas, apresentada na tabela 1. Comparando a coorte 1 com as coortes 2 e 3, nota-se um aumento nas porcentagens dos diagnósticos de AIDS: em mulheres, indivíduos heterossexuais e nas faixas etárias maiores que 30 anos. Em relação à escolaridade, a maioria dos diagnósticos ocorreu em indivíduos com escolaridade inferior a 8 anos. Houve uma redução da categoria de transmissão por uso de drogas injetáveis entre os pacientes notificados, o que pode ser visualizado nas coortes 2 e 3. Quando se avalia a variável óbito, observa-se uma queda drástica no número de mortes na coorte 3, quando comparada com as outras duas coortes. Nas três coortes, a maioria dos indivíduos notificados apresentavam contagem de células CD4 menor que 350 cel./mm<sup>3</sup> no momento do diagnóstico. Observa-se grande percentual de perdas de informações nos três períodos de estudo, como mostra a tabela 1. Devido à coorte 1 apresentar 96% de falta de informação da variável CD4, compatível com a realidade dos serviços na primeira década da epidemia, optou-se por não incluí-la no ajuste do modelo da coorte 1, e também considerá-la com uma variável contínua nos ajustes dos modelos para maximizar a informação.



Tabela 1: Número e proporções de indivíduos diagnosticados com AIDS de acordo com as variáveis nas três coortes estudadas, Campinas, São Paulo.

Variável	Coorte 1		Coorte 2		Coorte 3	
	N (%)	Valor p	N (%)	Valor p	N (%)	Valor p
<b>Sexo</b>						
Masculino	245(85,7)	<0,01*	1006(69,1)	<0,001*	831(61,9)	<0,01*
Feminino	41(14,3)		450(30,9)		511(38,1)	
<b>Escolaridade**</b>						
< 8 anos	81(50,6)	0,87	564(66,2)	<0,01*	484 (54,3)	0,01*
8 ≥ anos	79 (49,4)		288 (33,8)		407 (45,7)	
<b>Idade</b>						
< 30 anos	142 (49,6)	<0,01*	466 (32,0)	<0,01*	363 (27,0)	<0,01*
30 – 49 anos	133 (46,5)		897 (61,6)		850 (63,3)	
≥ 50 anos	11 (3,85)		93 (6,39)		129 (9,6)	
<b>Droga injetável**</b>						
Sim	100 (45,7)	0,20	391 (32,7)	<0,01*	242 (20,5)	<0,01*
Não	119 (54,3)		806 (67,3)		940 (79,5)	
<b>Homo./ Bissexual**</b>						
Sim	128 (55,9)	0,07	279 (23,1)	<0,01*	317 (25,5)	<0,01*
Não	101 (44,1)		930 (76,9)		924 (74,5)	
<b>Óbito</b>						
Sim	208 (72,7)	<0,01*	469 (32,2)	<0,01*	176 (13,1)	<0,01*
Não	78(27,3)		987 (67,8)		1116 (86,9)	
<b>CD4**</b>						
<b>cel./mm<sup>3</sup></b>						
≤ 350	8 (72,7)	<0,13	720 (93,6)	<0,01*	898 (82,4)	<0,01*
> 350	3(27,3)		49 (6,4)		192 (17,6)	

\* Estatisticamente significante para  $p < 0,05$ .

\*\* As variáveis Escolaridade, Droga Injetável, Homossexual/Bissexual e CD4 apresentaram, respectivamente, os seguintes números e porcentagens de falta de informação, Coorte 1: 126 (44%), 67 (23%), 57 (20%) e 275 (96%). Coorte 2: 604 (41%), 259 (18%), 247 (17%) e 687 (47%). Coorte 3: 451 (34%), 160 (12%), 101 (7%) e 252 (19%).

A Tabela 2 apresenta os resultados do ajuste do modelo semi-paramétrico com as estimativas posteriores (média, desvio padrão). A Tabela 3 mostra as razões de riscos associadas às variáveis do estudo e seus respectivos intervalos de credibilidade. Na coorte 1, os efeitos das variáveis sexo, escolaridade e categoria de transmissão por uso de drogas injetáveis não foram significativos, pois a média posteriori igual a zero se encontra no intervalo de credibilidade de 95%. Na coorte 2 e coorte 3, a variável categoria de transmissão por uso drogas injetáveis não foi significativamente associada ao risco de morrer. Na coorte 3, também se observa que os efeitos das variáveis idade, na faixa de 30 a 49 anos, escolaridade e contagem de CD4 não foram significativos. O risco de óbito, na coorte 3, para homens foi de 1,91 (IC 95%: 1,36 - 2,73) vezes maior do que para as mulheres. Para os maiores de 49 anos, o risco de óbito foi de 2,59 (IC 95%: 1,58 - 4,24) vezes o risco em menores que 30 anos. Observa-se que em todas as coortes ter idade maior que 49 anos foi fator de risco de óbito. Para aqueles com categoria de transmissão homossexuais/bissexual, o risco de óbito foi 0,48 (IC 95% 0,31 - 0,72) vezes (ou 52% menor) que o risco entre os da categoria de transmissão heterossexuais. Os mapas apresentados nas Figuras 1 e 2 (A, B e C), mostram o risco de óbito por setor censitário na cidade de Campinas, indicando que quanto maior o valor na escala, maior foi o risco de óbito dos indivíduos notificados por AIDS, nos períodos de estudo. Na Figura 1 foi ajustado o modelo apenas com o intercepto e na Figura 2 o modelo com as covariáveis, nas três coortes. Observa-se na coorte 1, que as regiões sudoeste, sul e central da cidade apresentaram maiores riscos (Figuras 1A e 2A). Na coorte 2 (Figuras 1B e 2B), as regiões sudoeste, noroeste, sul e central exibiram maiores riscos de óbito. Por fim, na coorte 3 (Figuras 1C e 2C), o risco de óbito foi maior entre as regiões norte, noroeste, parte da sudoeste e na região central. Os mapas da Figura 2 apresentam algumas diferenças com as da Figura 1, indicando a influência das covariáveis para o aumento do risco de óbito em certas regiões. Os mapas apresentam diferenças na distribuição espacial do risco de morrer, porém predomina o maior risco em áreas com indicadores socioambientais precários como mostra a Figura 1D. Na Figura 3 apresentam-se os gráficos de risco basal do

modelo ajustado, em cada umas das coortes. Destaca-se que o risco de óbito diminui ao longo tempo, nas coortes 2 e 3, e se torna próximo de zero em torno de 2000 dias de seguimento.

Tabela 2: Estimativas posteriores (Média, desvio padrão (DP) e intervalo de credibilidade de 95% - IC) dos parâmetros do modelo semi-paramétrico para as variáveis, nas coortes 1, 2 e 3.

Variável	Coorte 1		Coorte 2		Coorte 3	
	Média (DP)	IC	Média (DP)	IC	Média (DP)	IC
Intercepto	-8,78 (-0,44)	- 9,71; -7,99	-10,9 (-0,26)	-11,46; -10,43	-12,01 (-0,32)	-12,68; -11,99
Sexo						
Masculino	0,23 (-0,21)	-0,18; 0,66	0,444 (-0,12)	0,22; 0,67	0,65 (-0,18)	0,30; 1,02
Feminino						
Escolaridade						
<8 anos	0,01 (-0,16)	-0,31; 0,32	-0,34 (-0,10)	-0,53; -0,14	0,18 (-0,15)	-0,12; 0,49
≥8 anos						
Idade						
≤30 anos						
30 – 49 anos	0,78 (-0,15)	0,49; 1,08	0,25 (-0,11)	0,04; 0,46	(0,34) -0,20	-0,04; 0,74
≥50 anos	0,84 (-0,35)	0,12; 1,48	0,92 (-0,18)	0,56; 1,25	0,95 (-0,25)	0,46; 1,44
Droga injetável						
Sim	-0,31 (-0,16)	-0,63; 0,00	0,09 (-0,11)	-0,11; 0,30	0,01 (-0,19)	-0,38; 0,38
Não						
Homo./Bi.						
Sim	-0,72 (-0,17)	-1,05; -0,38	-0,57 (-0,13)	-0,84; -0,32	-0,72 (-0,22)	-1,16; -0,30
Não						
CD4*			-0,006 (0,001)	-0,007; -0,005	-0,001 (0,003)	-0,004; 0,004

\* contagem de células CD4 cel./mm<sup>3</sup> variável contínua.

Tabela 3: Razão de risco de óbito e seus respectivos intervalos de credibilidade de 95% (IC-RR) obtidos do ajuste do modelo semi-paramétrico, para os indivíduos diagnosticados com AIDS, nas coortes 1, 2 e 3, Campinas, SP.

Variável	Coorte 1		Coorte 2		Coorte 3	
	RR	IC-RR	RR	IC-RR	RR	IC-RR
<b>Sexo</b>						
Masculino	1,26	(0,84; 1,94)	1,55	(1,24; 1,95)	1,91	(1,36; 2,73)
Feminino	-	-	-	-	-	-
<b>Escolaridade</b>						
< 8 anos	1,01	(0,74; 1,37)	0,71	(0,59; 0,87)	1,20	(0,88; 1,63)
≥ 8 anos	-	-	-	-	-	-
<b>Idade</b>						
< 30 anos	-	-	-	-	-	-
30 – 49 anos	2,19	(1,62; 2,95)	1,28	(1,04; 1,58)	1,41	(0,96; 2,10)
≥ 50 anos	2,32	(1,13; 4,39)	2,50	(1,76; 3,50)	2,59	(1,58; 4,29)
<b>Droga</b>						
<b>injetável</b>						
Sim	0,73	(0,53; 1,00)	1,10	(0,89; 1,35)	1,01	(0,68; 1,46)
Não	-	-	-	-	-	-
<b>Homo./Bi.</b>						
Sim	0,49	(0,35; 0,68)	0,56	(0,43; 0,73)	0,49	(0,31; 0,74)
Não	-	-	-	-	-	-
<b>CD4</b>						
cel./mm <sup>3</sup>	-	-	0,994	(0,992; 0,995)	1,00	(0,99; 1,01)

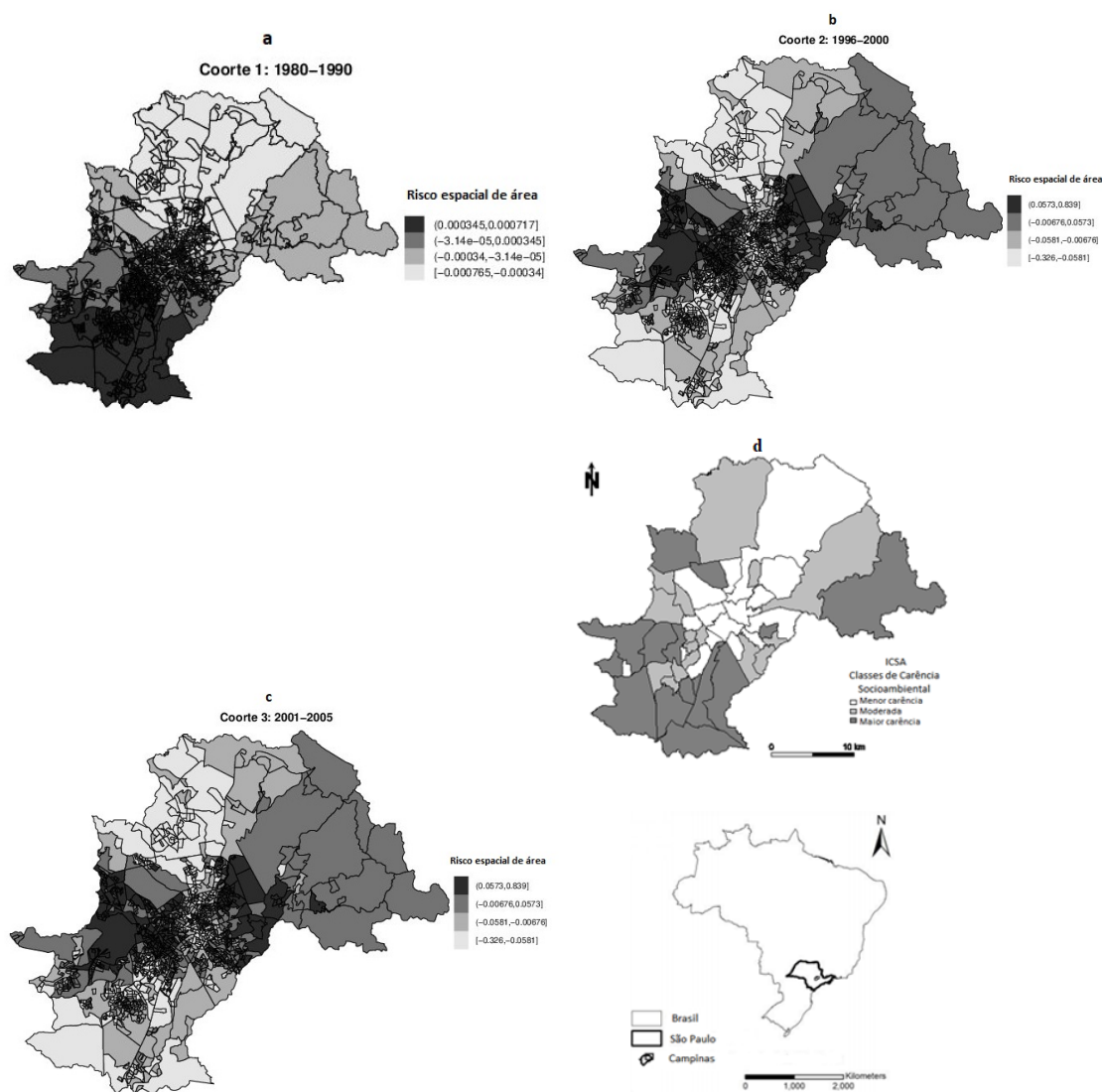


Figura 1: Mapas dos efeitos espaciais de área para os setores censitários (Médias posteriori), modelos ajustados apenas com o intercepto, para a mortalidade por AIDS nas Coortes estudadas (Figuras A, B e C) e Mapa de classes de Carência Socioambiental (Figura D\*) da cidade de Campinas-SP.

\*Figura d foi obtida do artigo Costa JV, et al.<sup>16</sup>

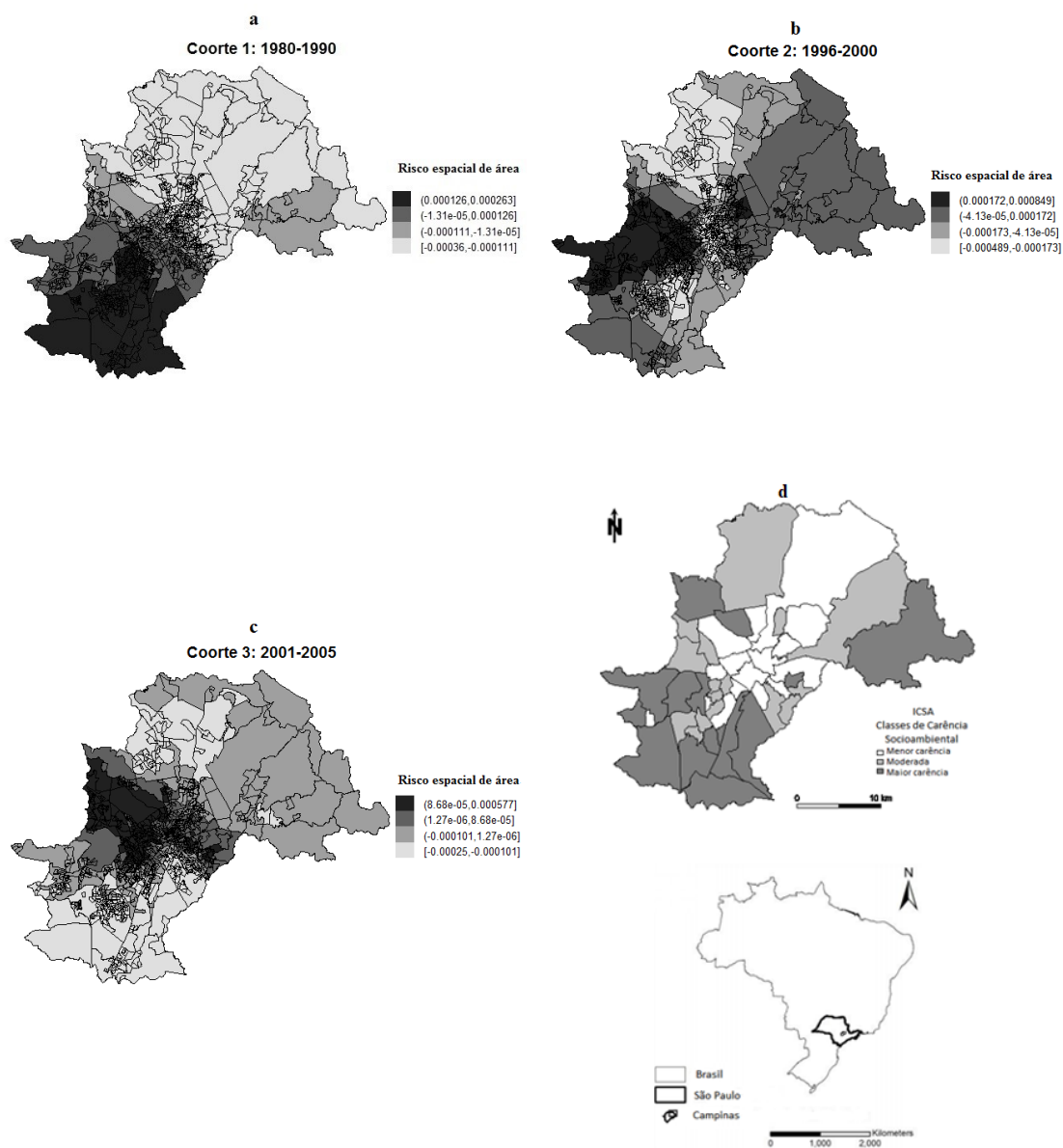


Figura 2: Mapas dos efeitos espaciais de área para os setores censitários (Médias posteriori), modelos ajustados com as covariáveis, para a mortalidade por AIDS nas Coortes estudadas (Figuras A, B e C) e Mapa de classes de Carência Socioambiental (Figura D\*) da cidade de Campinas-SP.

\*Figura d foi obtida do artigo Costa JV, et al.<sup>16</sup>

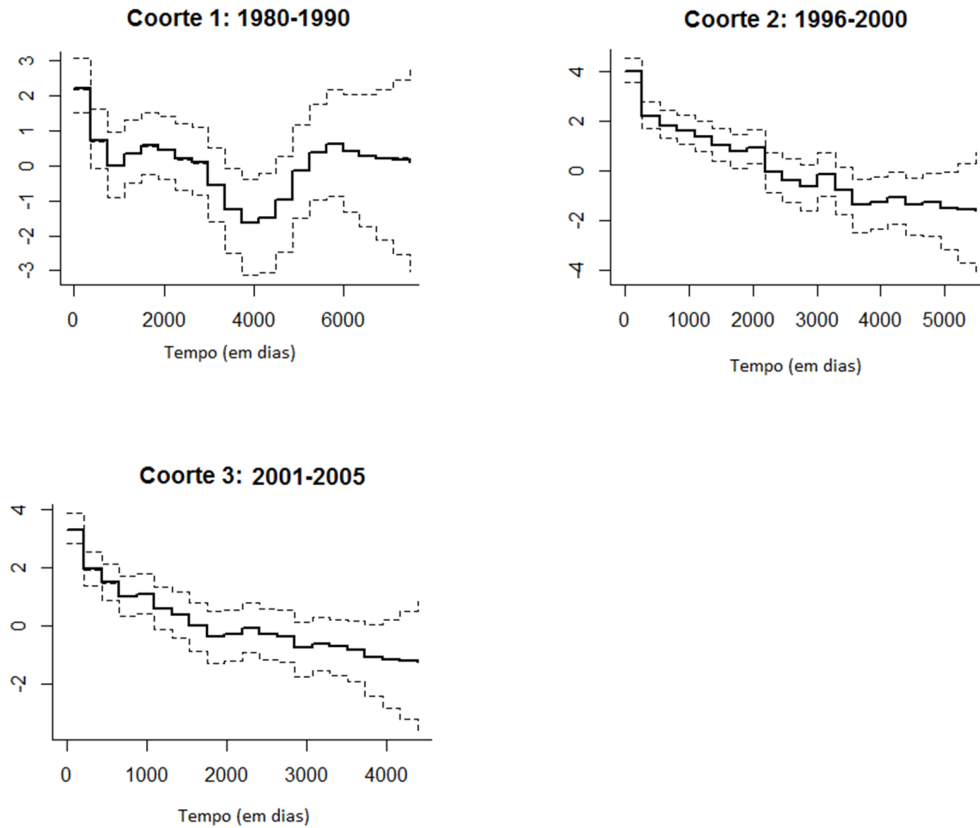


Figura 3: Média posteriori (linha sólida) e intervalo de credibilidade com 95% (linha tracejada) para o logaritmo do risco basal do modelo ajustado para a mortalidade.

## DISCUSSÃO

Este estudo evidencia a queda crescente do risco de morrer entre os pacientes notificados com AIDS e moradores do município de Campinas em três períodos. A mortalidade entre os pacientes diminuiu ao longo do tempo, como também tem sido apontado em outros estudos no Brasil e outras partes do mundo.<sup>3,4,6,17</sup> Com o acesso à TARV a partir de 1996 devido à disponibilização pelo Ministério da Saúde, obteve-se o controle mais efetivo da proliferação viral e da recuperação do sistema imune dos pacientes. Além do impacto na mortalidade, a morbidade dos pacientes com infecção pelo HIV tem diminuído progressivamente, elevando a sobrevida e a qua-

lidade de vida dos pacientes infectados. Contudo, continua sendo uma doença com grande impacto na morbidade e mortalidade e exige ações de prevenção e assistência contínuas. Os estimadores das análises nas coortes estudadas identificam os homens, os maiores de 50 anos e os com categoria de transmissão heterossexuais como os de maior risco de morrer nos períodos de estudo em Campinas. Vários autores têm associado a maior sobrevida dos pacientes com AIDS com a idade ao diagnóstico, sexo, situação socioeconômica e categoria de exposição como observado neste estudo em Campinas.<sup>8,17,18,19</sup> As condições socioeconômicas dos pacientes podem influenciar o risco de morrer devido às diferenças no acesso ao seguimento clínico, prevenção de complicações e na adesão aos antirretrovirais, as quais são menores entre as populações mais pobres.<sup>4,11,17,18</sup> Os mapas apresentados mostram piores estimativas de sobrevida nas regiões sul e sudoeste do município principalmente nas primeiras coortes, restando a região sudoeste como de maior risco na coorte 2001-2005, em todos os mapas, regiões coincidentes com áreas de elevada carência socioambiental.<sup>16</sup> A distribuição espacial do risco entre os setores censitários em Campinas mostra diferenças entre as coortes. Embora em escalas decrescentes de risco, os mapas indicam que as regiões mais pobres da cidade (noroeste, sudoeste) sempre foram as com maior mortalidade associada à AIDS. As possíveis justificativas são as precárias condições de vida, de acesso aos serviços ou ao maior percentual de uso de drogas nas primeiras décadas da epidemia. O “envelhecimento” da epidemia e o pior prognóstico em maiores de 50 anos têm sido também identificado por outros autores.<sup>20</sup> Ressalta-se a necessidade de reconhecer a vulnerabilidade à infecção nesta faixa de idade, considerando-se a menor percepção de risco, a falta de conhecimento e de utilização de métodos de prevenção, além do diagnóstico tardio nesta faixa etária.<sup>19</sup> Algumas variáveis clínicas, não abordadas neste estudo, também têm sido apontadas como preditores de sobrevida de pacientes com AIDS, como as doenças oportunistas, comorbidades, acesso e adesão aos esquemas terapêuticos e a serviços especializados.<sup>3,4,6,17</sup> A maioria dos indivíduos com contagem  $CD4 \leq 350$  cel./mm<sup>3</sup> no momento do diagnóstico sugere o diagnóstico tardio da doença. A associação negativa da contagem de CD4 na



primeira consulta (variável contínua) com a mortalidade na coorte 2 reforça este parâmetro laboratorial como preditor de evolução clínica dos casos. Entretanto, na coorte 3, outras variáveis deslocaram o CD4 para explicar o risco de morrer, entre elas a idade e o sexo. O acesso à medicação antirretroviral combinada nos serviços de referência na cidade possivelmente diminuiu a influência deste indicador laboratorial como preditor de morte. Campinas conta com dois centros municipais de referência para DST/AIDS, um centro de orientação e apoio sorológicos, além de ambulatórios de referência do Hospital das Clínicas da Unicamp e no Hospital Celso Pierro da PUC de Campinas, com cobertura ampla de indivíduos com HIV/AIDS do município e região metropolitana. Algumas limitações do estudo podem ser apontadas como as decorrentes da utilização de dados secundários para a análise. A notificação do caso nem sempre ocorre no momento do diagnóstico da AIDS, muitos são detectados tardiamente, outros no momento de intercorrências clínicas ou do início do tratamento. A exclusão de pacientes notificados como AIDS na data do óbito procurou reduzir estas distorções. A subnotificação de casos de AIDS e óbitos pode ocorrer devido à falta de diagnóstico, bem como ao mau preenchimento de atestados de óbitos. Além disso, as fichas de notificação nem sempre são bem preenchidas, restando informações às vezes imprecisas. Porém trata-se de um banco de dados com informações valiosas sobre as tendências da doença no território e no tempo. Por outro lado, o longo período de seguimento das três coortes implica na comparação de coortes em diferentes contextos de risco, opções terapêuticas e de prevenção, com impacto no prognóstico dos pacientes. Os períodos de estudo certamente são heterogêneos, pois resultam de conjunturas epidemiológicas distintas.

## **CONCLUSÕES**

O estudo contribui para possíveis ações e monitoramento dos serviços de saúde do município de Campinas em relação aos grupos e às regiões da cidade de maior risco de morrer, entre elas, as de maior carência e piores condições socioambientais. Ações de prevenção e controle da doença e particularmente medidas para aumentar a adesão

dos pacientes ao seguimento clínico e ao uso correto da medicação disponível podem ser melhor direcionadas. Mesmo diante da diminuição do risco de morrer por AIDS nos últimos anos, após o acesso à TARV, a maior incidência se concentra entre os homens, mais idosos, categoria de transmissão heterossexuais e entre os mais pobres nas três coortes estudadas. Os mapas obtidos e os fatores de risco orientam possíveis ações e monitoramento da doença no município.

## REFERÊNCIAS

1. Fonseca MGP, Bastos FI. Twenty-five years of the AIDS epidemic in Brazil: principal epidemiological findings, 1980-2005. *Cad saude publica / Minist da Saúde, Fund Oswaldo Cruz, Esc Nac Saúde Pública*. 2007. doi:10.1590/S0102-311X2007001500002.
2. Fordyce EJ, Singh TP, Nash D, Gallagher B, Forlenza S. Survival rates in NYC in the era of combination ART. *J Acquir Immune Defic Syndr*. 2002; 30(1):111-118.
3. Lima VD, Hogg RS, Harrigan PR, et al. Continued improvement in survival among HIV-infected individuals with newer forms of highly active antiretroviral therapy. *AIDS*. 2007; 21(6):685-692. doi:10.1097/QAD.0b013e32802ef30c.
4. Schneider MF, Gange SJ, Williams CM, et al. Patterns of the hazard of death after AIDS through the evolution of antiretroviral therapy: 1984-2004. *AIDS*. 2005. doi:10.1097/01.AIDS.0000189864.90053.22.
5. Chequer P, Hearst N, Hudes ES, et al. Determinants of survival in adult Brazilian AIDS patients, 1982-1989. The Brazilian State AIDS Program Co-ordinators. *AIDS, London*. 1992. doi:10.1097/00002030-199205000-00007.
6. Guibu IA, Barros MB de A, Donalisio MR, Tayra Â, Alves MCGP. Survival of AIDS patients in the Southeast and South of Brazil: analysis of the 1998-1999 cohort. *Cad saude publica / Minist da Saúde, Fund Oswaldo Cruz, Esc Nac Saúde Pública*. 2011;27 Suppl 1:S79-S92.
7. Tancredi MV, Waldman EA. Survival of AIDS patients in Sao Paulo-Brazil in the pre- and post-HAART eras: a cohort study. *BMC Infect Dis*. 2014. doi:10.1186/s12879-014-0599-8.

8. Teixeira TR de A, Gracie R, Malta MS, et al. Social geography of AIDS in Brazil: identifying patterns of regional inequalities. *Cad Saude Publica*. 2014; 30(2):259-271. doi:10.1590/0102-311X00051313.
9. Sousa AIA de, Pinto VL, Sousa AIA de, Pinto VL. Análise espacial e temporal dos casos de AIDS no Brasil em 1996-2011: áreas de risco aumentado ao longo do tempo\*. *Epidemiol e Serviços Saúde*. 2016; 25(3):467-476. doi:10.5123/S1679-49742016000300003.
10. Santos NJS, Tayra A, Silva SR, Buchalla CM, Laurenti R. A AIDS no Estado de São Paulo: as mudanças no perfil da epidemia e perspectivas da vigilância epidemiológica. *Rev Bras Epidemiol*. 2002; 5(3):286-310. doi:10.1590/S1415-790X2002000300007.
11. Pellini ACG, Bellini JH, Cavalin RF, Neto FC, Zanetta DMT. Distribuição espacial das notificações de AIDS em mulheres no município de São Paulo nos períodos de 1999-2001 e 2009-2011: uma análise sob a ótica da desigualdade socioespacial. *Rev Espaço e Geogr*. 2015; 18(2).
12. Martino S, Akerkar R, Havard R. Approximate Bayesian Inference for Survival Models. *Scand J Stat*. 2011; 38(3):514-528. doi:10.1111/j.1467-9469.2010.00715.x.
13. R Development Core Team. R: A language and environment for statistical computing. 2017. <http://www.r-project.org/>. R Found Stat Comput Vienna, Austria.
14. Besag J, Kooperberg C. On conditional and intrinsic autoregressions. *Biometrika*. 1995; 82(4):733-746.
15. Rue H, Martino S, Chopin N. Approximate Bayesian inference for latent Gaussian models by using integrated nested Laplace approximations. *J R Stat Soc Ser B Stat Methodol*. 2009;71(2):319-392.

16. Costa JV, Donalisio MR, Silveira LV de A. Spatial distribution of dengue incidence and socio-environmental conditions in Campinas, São Paulo State, Brazil, 2007. *Cad Saude Publica*. 2013; 29(8):1522-1532. doi:10.1590/0102-311X00110912.
17. Burch LS, Smith CJ, Phillips AN, Johnson MA, Lampe FC. Socioeconomic status and response to antiretroviral therapy in high-income countries. *AIDS*. 2016; 30(8):1147-1161. doi:10.1097/QAD.0000000000001068.
18. Arnold M, Hsu L, Pipkin S, McFarland W, Rutherford GW. Race, place and AIDS: The role of socioeconomic context on racial disparities in treatment and survival in San Francisco. *Soc Sci Med*. 2009. doi:10.1016/j.socscimed.2009.04.019.
19. Barboza R. Homens idosos e o HIV/AIDS no campo da Saúde Coletiva: vulnerabilidades e desafios na quarta década da epidemia . *BIS Bol do Inst Saúde* . 2012;14:81-89.
20. Sousa ACA, Suassuna DSB, Costa SML. Perfil Clínico Epidemiológico de Idosos com AIDS/DST; DST-J Bras. *Doenças Sex Transm*. 2009.

## Artigo 2

**Análise de sobrevida com efeito espacial de indivíduos diagnosticados com AIDS no município de Campinas em três coortes: 1980-1990, 1996-2000 e 2001-2005.**

Thiago Santos Mota<sup>1</sup>, Ana Carolina Nunes Mafra<sup>2</sup>, Maria Rita Donalisio<sup>3</sup>, Liciania Vaz de Arruda Silveira<sup>4</sup>

Departamento de Bioestatística. Universidade Estadual Paulista “Júlio de Mesquita Filho” – UNESP. SP, Brasil<sup>1,4</sup>

Instituto Israelita de Ensino e Pesquisa, Hospital Israelita Albert Einstein. SP, Brasil<sup>2</sup>

Departamento de Saúde Coletiva. Faculdade de Ciências Médicas. Universidade Estadual de Campinas – UNICAMP. SP, Brasil<sup>3</sup>

## RESUMO

O objetivo desse estudo foi analisar a sobrevida e o risco de óbito, em indivíduos maiores que 13 anos com AIDS no município de Campinas-SP, em função de variáveis sociodemográficas, comportamentais e clínicas, considerando a sua localização espacial (latitude e longitude). Os dados foram obtidos do Sistema de Informação de Agravos de Notificação (SINAN) e dos óbitos que constam no Sistema de Informação sobre Mortalidade (SIM) de moradores de Campinas, no período de 1980 a 2014. As análises de sobrevida foram realizadas em três coortes que se referem aos anos de diagnóstico/notificação (1980-1990, 1996-2000 e 2001-2005) com seguimento retrospectivo de 10 anos. Modelos de riscos proporcionais com efeitos espaciais aditivos, uma extensão do modelo de Cox incorporada nos modelos aditivos generalizados, foram ajustados, possibilitando obter os riscos espaciais de óbito associado à AIDS no município. As análises mostram aumento da sobrevida ao longo dos anos e diferentes variáveis associadas à AIDS nos períodos analisados. Nos pacientes da coorte 1, idade maior que 30 anos foi associadas a maior mortalidade; na coortes 2, idade maior ou igual a 50 anos, ser usuário de drogas injetáveis, ter tuberculose e apresentar contagem de células CD4 menor ou igual a 200 cél./mm<sup>3</sup> de sangue (imunossupressão grave); e na coorte 3, o modelo foi estratificado pela variável imunossupressão grave, por esta variável ferir a pré-suposição de riscos proporcionais, indicando que ser do sexo masculino e possuir baixa escolaridade estão associados a menor sobrevida. A variação espacial para o risco de morte foi significativa para a coorte 1 e não significativo para as coortes 2 e 3, indicado que a localização espacial nestas duas coortes não foi um fator principal para o risco de óbito entre os pacientes. O estudo mostra contextos diferentes nas diversas etapas da epidemia em Campinas indicando que, mesmo em uma cidade com o programa de IST-AIDS bem consolidado, é necessário acompanhamento contínuo para conhecer os fatores associados à menor sobrevida entre os pacientes, devido às mudanças da epidemia ao longo do tempo.

**Palavras-chave:** AIDS. Sobrevida. Distribuição espacial.

## INTRODUÇÃO

No mundo 22 milhões de pessoas não têm acesso ao tratamento para a infecção pelo HIV, mesmo diante de metas otimistas de fim da epidemia propostas pela UNAIDS no documento 90/90/90<sup>1</sup>, que preconiza 90% de diagnóstico, de tratamento e de supressão viral em 2020. Estes percentuais continuam sendo um desafio para o Brasil e para o mundo, dada as discrepâncias do acesso ao tratamento e ao seguimento clínico associadas às desigualdades sociais.<sup>1,2,3,4</sup> O Ministério da Saúde do Brasil tem notificado em média 40 mil casos de AIDS por ano desde 2013, em diferentes proporções conforme a região do país.<sup>5</sup> Em Campinas, estado de São Paulo, são registrados em média 250 a 300 casos novos por ano pelo Sistema de Vigilância Epidemiológica. Mudanças no perfil da epidemia têm sido registradas desde o início de seu monitoramento, trazendo desafios aos serviços de saúde para seu enfrentamento. A análise contínua da incidência e letalidade da infecção pelo HIV e fatores associados pode trazer informações sobre os grupos mais afetados e medidas de prevenção e controle mais precisas.<sup>6</sup> Estimadores de sobrevivência de pacientes com AIDS no Brasil e sua distribuição espacial têm sido pouco abordados na literatura.<sup>7</sup> Embora não reflitam o padrão espacial de acesso dos indivíduos aos serviços assistenciais e de prevenção, indiretamente refletem condições socioeconômicas relevantes para o prognóstico da doença.<sup>8</sup>

Apesar da redução drástica da mortalidade por AIDS no Brasil e em várias partes do mundo, em decorrência da utilização da terapia antirretroviral (TARV) e de outras estratégias assistenciais e de prevenção, tem se observado no Brasil tendência de aumento da incidência e da mortalidade de pacientes com infecção pelo HIV/AIDS nos últimos anos.<sup>6</sup>

No Brasil, o acesso à medicação foi disponibilizado pelo Ministério da Saúde desde 1996, ocasionando impacto na sobrevivência dos pacientes como apontam vários estudos.<sup>3,9,10</sup> Além disso, observou-se uma redução no número de internações de pessoas com AIDS, diminuição de infecções oportunistas e de doenças crônicas.<sup>3,11,12,13</sup>



O objetivo deste estudo foi analisar a sobrevida e o risco de óbitos de indivíduos com infecção do HIV/AIDS com o efeito espacial e variáveis associadas, em três coortes de pacientes notificados, considerando-se diferentes períodos da epidemia, em uma metrópole do estado de São Paulo.

## MATERIAIS E MÉTODOS

Foram constituídas três coortes retrospectivas de pacientes notificados por AIDS, para se analisar a sobrevida e o risco espacial de óbito no período de 10 anos, de indivíduos maiores de 13 anos, moradores do município de Campinas, a nordeste do estado de São Paulo (aproximadamente 1 milhão de habitantes). As informações foram obtidas do Sistema de Informação de Agravos de Notificação (SINAN) e os óbitos do Sistema de Informação sobre Mortalidade (SIM) de moradores de Campinas (1980 a 2014). Uma coorte foi composta por indivíduos notificados, de 1980 a 1990 (coorte 1 – a primeira década da epidemia com restritas opções terapêuticas), outra com as notificações de 1996 a 2000 (coorte 2 – após a disponibilização de terapia antirretroviral de alto impacto), e uma terceira com pacientes registrados no período de 2001 a 2005 (coorte 3 – seguimento de pacientes até antes do início do tratamento para todos os infectados). Não foram incluídos pacientes que não constavam na lista de notificação do SINAN.

Após checagem de duplicações e inconsistência dos bancos, as variáveis utilizadas no estudo foram: sociodemográficas (sexo, escolaridade, idade com que o indivíduo foi diagnosticado com AIDS), variáveis comportamentais – uso de droga injetáveis, preferência sexual (homens que fazem sexo com homens, HSH, bissexual ou heterossexual), condições clínicas associadas à sobrevida – tuberculose (todas as formas clínicas)<sup>14</sup>, câncer<sup>15</sup> (pelo menos um dos seguintes diagnósticos: linfoma não Hodgkin, linfoma primário do cérebro, câncer de colo uterino ou Sarcoma de Kaposi), marcadores laboratoriais – imunossupressão grave (contagem de células T CD4+ no momento do diagnóstico categorizada como  $\leq$  ou  $>$  que 200 cél./mm<sup>3</sup>) e as localizações das residências georreferenciadas (latitude e longitude). Para realizar a mo-

delagem utilizou-se o modelo de riscos proporcionais<sup>16</sup> e uma extensão desse modelo incorporada nos modelos lineares aditivos generalizados (GAM)<sup>17,18</sup>, que permitiu a obtenção do risco espacial controlado por fatores individuais de risco. Considerando como variável resposta os tempos até o óbito, por qualquer causa, dos  $n$  pacientes notificados por AIDS em cada coorte denotados por  $t_1 < \dots < t_k$ , com  $k < n$ ,  $\mathbf{X}_i$  o vetor de covariáveis para o  $i$ -ésimo indivíduo e a função de suavização  $f(s_i)$ , com  $i = 1, \dots, n$ , que incorpora a localização espacial (latitude e longitude) do  $i$ -ésimo indivíduo na região que compreende a cidade de Campinas. O modelo de riscos proporcionais com efeito espacial aditivo, no qual foi ajustado para o  $i$ -ésimo indivíduo, é apresentado da seguinte forma:

$$\begin{aligned} \lambda(t | \mathbf{X}_i, s_i) &= \lambda_0(t) \exp \{ \mathbf{X}_i' \boldsymbol{\beta} + f(s_i) \} \\ &= \lambda_0(t) \exp \{ \beta_1 \cdot \text{sexo} + \beta_2 \cdot \text{idade} + \beta_3 \cdot \text{escolaridade} + \beta_4 \cdot \text{droga injetável} \\ &\quad + \beta_5 \cdot \text{HSH/bissexual} + \beta_6 \cdot \text{câncer} + \beta_7 \cdot \text{tuberculose} \\ &\quad + \beta_8 \cdot \text{imunossupressão} + f(s_i) \}. \end{aligned}$$

Devido ao número reduzido de observações na variável imunossupressão, na coorte 1, esta variável não foi ajustado no modelo. Para a modelagem do efeito espacial  $f(s_i)$  utilizou-se a função de *spline* bivariado incorporada linearmente como uma covariável do modelo. Todos os ajustes do modelo GAM e os mapas de riscos de óbito foram obtidos por meio de funções do pacote *mgcv*<sup>19</sup>. As funções desse pacote não delimitam os riscos nos limites do município, o que foi necessário elaborar uma função para facilitar as interpretações dos mapas. Todas as análises estatísticas foram realizadas no programa R<sup>20</sup> e o mapa da cidade de Campinas-SP foi obtido no site do IBGE, utilizando a malha do setor censitário urbano do ano 2000, como um ano intermediário do período de estudo. Os setores censitários do *shapefile* não foram considerados na análise, afim de obter uma medida de risco que possa variar continuamente dentro dos limites da cidade de Campinas.

## RESULTADOS

Na coorte 1 foram analisados 286 indivíduos, sendo 245 (85,66%) homens; na coorte 2, os dados correspondiam a 1456 indivíduos, consistindo em 1006 (69,09%) homens e na coorte 3, observou-se 1342 indivíduos, sendo 831 (61,92%) homens. A tabela 1 apresenta as proporções de óbitos nas três coortes segundo as variáveis de estudo. Observa-se redução do risco de óbito em todas as categorias das coortes 1 para a 2, e desta para a coorte 3, e a letalidade diminuiu progressivamente nas três coortes de estudo. Porém, é possível notar um aumento nas porcentagens de diagnósticos de AIDS entre as mulheres, maior mortalidade em indivíduos: com escolaridade inferior a 8 anos, heterossexuais e do sexo masculino. As curvas de sobrevivência e o teste de *logrank* de cada variável, em cada coorte, são apresentadas nas figuras 1, 2 e 3. Na coorte 1 o tempo mediano de sobrevida foi inferior a 2000 dias, quando observado as curvas de sobrevivência dos indivíduos em todas as covariáveis, observa-se pior sobrevida dos indivíduos com idade superior a 30 anos. Nas coortes 2 e 3 (figuras 2 e 3) há um aumento na probabilidade de sobrevida dos indivíduos, superior a 50% em todas as variáveis analisadas no período.

Na figura 2 os gráficos das curvas de sobrevida, em que o teste de *logrank* foi significativo ao nível de 5% de significância, mostram que ser do sexo masculino, ser usuário de droga injetáveis, ter idade superior a 50 anos, baixa escolaridade, tuberculose ou imunossupressão grave, são indicadores de pior sobrevida dos indivíduos nesta coorte.

Na figura 3 é possível notar a mudança nas curvas de sobrevida da coorte 2 para coorte 3, com exceção da variável droga injetável, todas as variáveis do estudo, ao nível de 5% significância, apresentaram diferença estatística significativa quando comparada às curvas em cada variável. Além disso, observa-se, um aumento na probabilidade de sobrevida dos indivíduos.

No processo de modelagem, ajustamos o modelo de Cox sem o efeito espacial, com todas as covariáveis do estudo (sexo, escolaridade, idade, uso de droga injetável,

HSH/bissexual, câncer, tuberculose e imunossupressão). A variável imunossupressão não foi ajustada na coorte 1, devido às poucas observações dessa covariável (tabela 1). Em seguida realizou-se o teste de proporcionalidade dos riscos através dos resíduos de Schoenfeld e verificou-se que não houve problemas com a suposição de riscos proporcionais em nenhuma das covariáveis do modelo, nas coortes 1 e 2. Na coorte 3 a variável imunossupressão violou a suposição de riscos proporcionais ( $p = 0,00594$ ), através dessa informação e por se tratar de uma variável importante em dados de AIDS, optou-se em estratificar o modelo pela variável imunossupressão nesta coorte. A partir dessa análise preliminar, ajustou-se o modelo de Cox aditivo incluindo o efeito espacial, no qual foi significativo apenas na coorte 1, como apresentado na tabela 2. Fazendo a seleção das variáveis pelo método stepwise, no modelo apresentado na tabela 2, o modelo final para cada uma das coortes é apresentado na tabela 3. Observa-se na coorte 1, tabela 3, que a razão de risco de óbito para indivíduos na faixa etária 30 a 49 anos foi de aproximadamente 1,80 vezes maior do que para os com menos que 30 anos, e para a faixa etária maior ou igual a 50 anos foi de 2,86 vezes maior do que para os indivíduos com menos de 30 anos.

A análise indica haver variação espacial do risco de mortalidade em pacientes com AIDS apenas na coorte 1 (Tabela 3, Figura 4), sendo que a região sul da cidade nesse período apresenta os maiores riscos. Na coorte 2, observa-se que ter 50 anos ou mais, ser usuário de droga, ter tuberculose e possuir imunossupressão grave são considerados como fatores de risco de óbito (Tabela 3). Nesta coorte não houve variação espacial significativa do risco de óbito por AIDS ( $p = 0,32$ ). Na coorte 3, o modelo estratificado por imunossupressão, apresentou como fatores de risco baixa escolaridade e sexo masculino (Tabela 3). Nesta coorte, a variação espacial do risco também não foi significativa ( $p = 0,58$ ). Os gráficos de resíduos desses modelos ajustados são apresentados na figura 5, evidenciando alguns pontos com valores maiores que 2 e menores que -2. Os gráficos não apresentam nenhuma tendência, mas nas coortes 2 e 3 observamos uma separação bem clara dos indivíduos que falharam (vieram a óbito) dos indivíduos censurados, o que pode ser explicado pela

queda drástica da mortalidade nessas duas coortes.

Tabela 1: Característica dos indivíduos diagnosticados com AIDS nas três coortes estudadas em Campinas, São Paulo. Tamanho da amostra (n) e percentual (%), número de mortes (m) e percentual (%) de mortes e valor p para as comparações das proporções de mortes pelo teste Qui-Quadrado.

Características	Coorte 1			Coorte 2			Coorte 3		
	(n=286)			(n=1456)			(n=1342)		
	n (%)	m (%)	Valor p	n (%)	m (%)	Valor p	n(%)	m (%)	Valor p
Sexo									
Masculino	245 (85,66)	179 (73,1)	0,9	1006 (69,09)	362 (35,9)	<0,01	831 (61,92)	132 (15,9)	<0,01
Feminino	41 (14,34)	29 (70,7)		450 (30,90)	109 (24,1)		511 (38,07)	54 (10,6)	
Idade									
≤30 anos	142 (49,65)	92 (64,8)	0,01	466 (32,01)	126 (27,0)	<0,01	386 (28,76)	36 (9,3)	<0,01
30-49 anos	133 (46,50)	106 (79,7)		897 (61,61)	298 (33,2)		821 (61,18)	116 (14,1)	
≥50 anos	11 (3,85)	10 (90,9)		93 (6,39)	47 (50,5)		135 (10,06)	34 (25,2)	
Escolaridade *									
< 8 anos	81 (50,63)	61 (75,3)	0,33	564 (66,20)	169 (29,9)	<0,01	484 (54,32)	80 (16,5)	<0,01
≥ 8 anos	79 (49,38)	53 (67,1)		288 (33,80)	53 (18,4)		407 (45,68)	32 (7,9)	
HSH/Bissexual*									
Sim	128 (55,90)	85(66,4)	0,13	279 (23,08)	72 (25,8)	0,44	317 (25,54)	29 (9,1)	0,03
Não	101 (44,10)	77 (76,2)		930 (76,92)	264 (28,4)		924 (74,46)	129 (13,9)	
Droga injetável*									
Sim	119 (54,34)	87 (73,1)	0,4	391 (32,66)	142 (36,3)	<0,01	242 (20,47)	38 (15,7)	0,14
Não	100 (45,66)	67 (67,0)		806 (67,34)	208 (25,8)		940 (79,53)	112 (11,9)	
Câncer*									
Sim	33 (20,25)	23 (69,7)	0,92	42 (3,12)	16 (38,1)	0,42	22 (1,85)	9 (40,4)	<0,01
Não	130 (79,75)	87 (66,9)		1306 (96,88)	405 (31,0)		1167 (98,15)	163 (13,9)	
Tuberculose*									
Sim	35 (23,03)	22 (62,8)	0,42	219 (16,86)	90 (41,1)	<0,01	115 (10,84)	32 (27,8)	<0,01
Não	117 (76,97)	84 (71,8)		1080(83,14)	320 (29,62)		946 (89,16)	129 (13,6)	
Imunossupressão *									
≤200 cel./mm <sup>3</sup>	6 (54,55)	2 (33,3)	1	434 (56,44)	101 (23,3)	<0,01	510 (46,79)	72 (14,1)	<0,01
>200 cel./mm <sup>3</sup>	5 (45,45)	1 (20,0)		335 (43,56)	38 (11,3)		580 (53,21)	36 (6,2)	

\* As variáveis escolaridade, HSH/bissexual, droga injetável, câncer, tuberculose e imunossupressão grave apresentaram, respectivamente, os seguintes números e porcentagens de falta de informação, Coorte 1: 126 (44%), 57 (20%), 67 (23%), 123 (43%), 134(47%), 275 (96%); Coorte 2: 604 (41%), 247 (17%), 259 (18%), 108 (7%), 157 (11%), 687 (47%); Coorte 3: 451 (34%), 101 (7%), 160 (12%), 153 (11%), 281 (21%), 252 (19%).

Tabela 2: Razões de risco obtidas do ajuste do modelo de riscos proporcionais aditivo com todas as covariáveis, para os dados de AIDS, nas três coortes, com intervalo de confiança de 95%, Campinas, São Paulo.

Variáveis	Coorte 1			Coorte 2			Coorte 3		
	RR	IC95%	Valor p	RR	IC95%	Valor p	RR	IC95%	Valor p
Sexo									
Masculino	0,44	(0,17;1,13)	0,088	1,37	(0,73;2,55)	0,326	1,91	(0,97; 3,76)	0,06
Feminino									
Idade									
<30 anos									
30-49 anos	5,59	(2,64;11,83)	<0,001	1,33	(0,74;2,41)	0,344	1,07	(0,57;2,00)	0,834
≥50 anos	32,27	(2,87;362,36)	0,005	4,65	(1,72;12,55)	0,002	1,48	(0,63;3,46)	0,368
Escolaridade									
<8 anos	0,45	(0,18; 1,11)	0,081	1,28	(0,71;2,31)	0,402	1,74	(0,99;3,07)	0,053
≥8 anos									
HSH/bissexual									
Sim	1,11	(0,43;2,88)	0,828	0,59	(0,30;1,16)	0,124	0,97	(0,52;1,83)	0,92
Não									
Droga injetável									
Sim	1,76	(0,79;3,93)	0,1676	2,88	(1,62;5,12)	<0,001	1,5	(0,82;2,72)	0,188
Não									
Câncer									
Sim	0,85	(0,28;2,51)	0,7633	1,28	(0,30;5,55)	0,738	2,33	(0,54;10,04)	0,257
Não									
Tuberculose									
Sim	0,78	(0,29;2,14)	0,6351	1,76	(0,93;3,32)	0,081	1,44	(0,63;3,29)	0,3836
Não									
Imunossupressão									
≤200 cel./mm <sup>3</sup>				2,12	(1,23; 3,69)	0,007			
>200 cel./mm <sup>3</sup>									
Efeito espacial			0,0262			0,427			0,547

Tabela 3: Razões de risco obtidas do ajuste do modelo de riscos proporcionais aditivo após a seleção das covariáveis, para os dados de AIDS, nas três coortes, com intervalo de confiança (IC) de 95%, Campinas, São Paulo.

Variáveis	Coorte 1		
		Razão de Risco (IC)	Valor p
Idade	$\leq 30$ anos		
	30-49 anos	1,80 (1,35 ; 2,39)	<0,01
	$\geq 50$ anos	2,86 (1,47 ; 5,55)	<0,01
Efeito espacial			0,05
Variáveis	Coorte 2		
		Razão de Risco (IC)	Valor p
Idade	$\leq 30$ anos		
	30-49 anos	1,39 (0,89 ; 2,17)	0,15
	$\geq 50$ anos	2,89 (1,32 ; 6,32)	<0,01
Droga injetável	Sim	2,13 (1,43; 3,18)	<0,01
	Não		
Tuberculose	Sim	1,79 (1,08; 2,97)	<0,01
	Não		
Imunossupressão Grave	$\leq 200$ cél./mm <sup>3</sup>	2,35 (1,54 ; 3,60)	<0,01
	$> 200$ cél./mm <sup>3</sup>		
Efeito espacial			0,32
Variáveis	Coorte 3		
		Razão de Risco (IC)	Valor p
Escolaridade	$\leq 8$ anos	2,07 (1,27 ; 3,37)	<0,01
	$> 8$ anos		
Sexo	Masculino	1,70 (1,01; 2,86)	0,04
	Feminino		
Efeito espacial			0,58

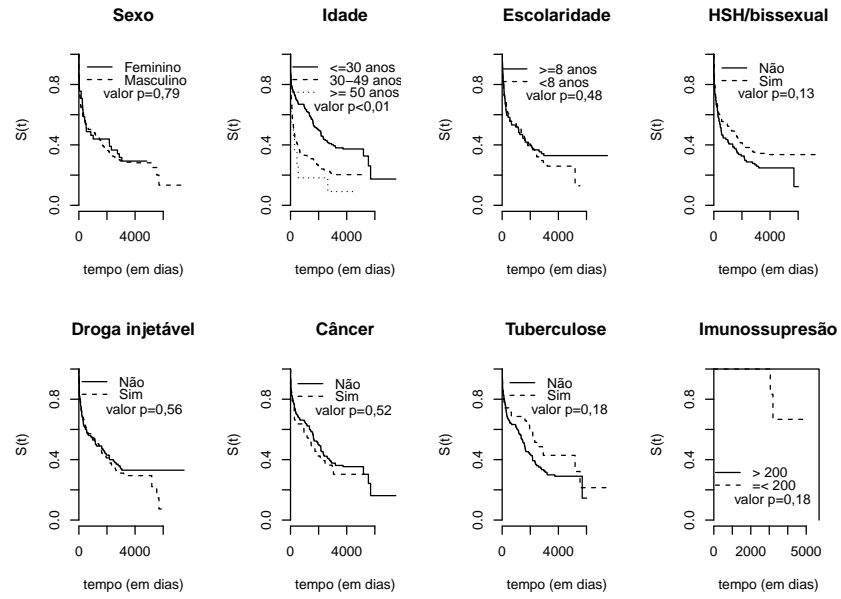


Figura 1: Curvas de sobrevivência em relação a cada variável analisada e valor p para o teste de logrank, na coorte 1 (1980-1990), Campinas, São Paulo.

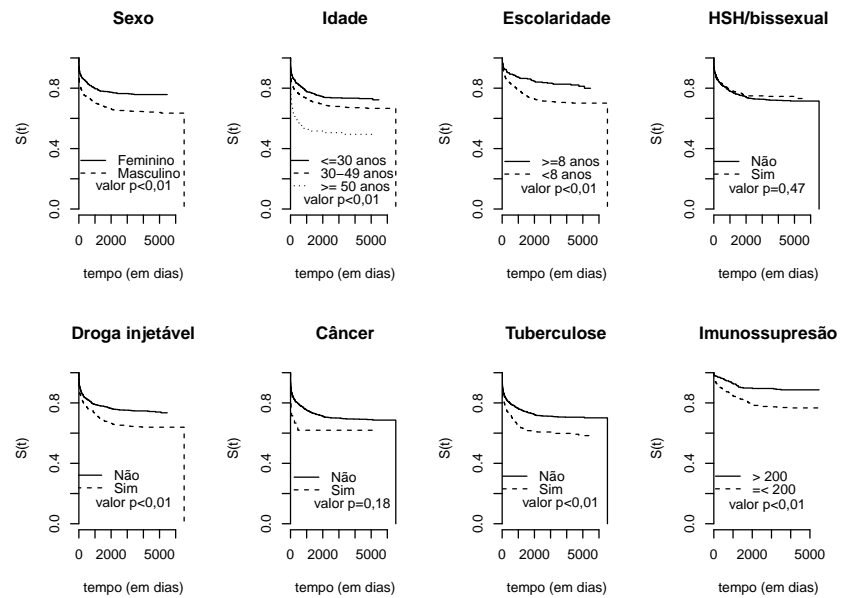


Figura 2: Curvas de sobrevivência em relação a cada variável analisada e valor p para o teste de logrank, na coorte 2 (1996-2000), Campinas, São Paulo.



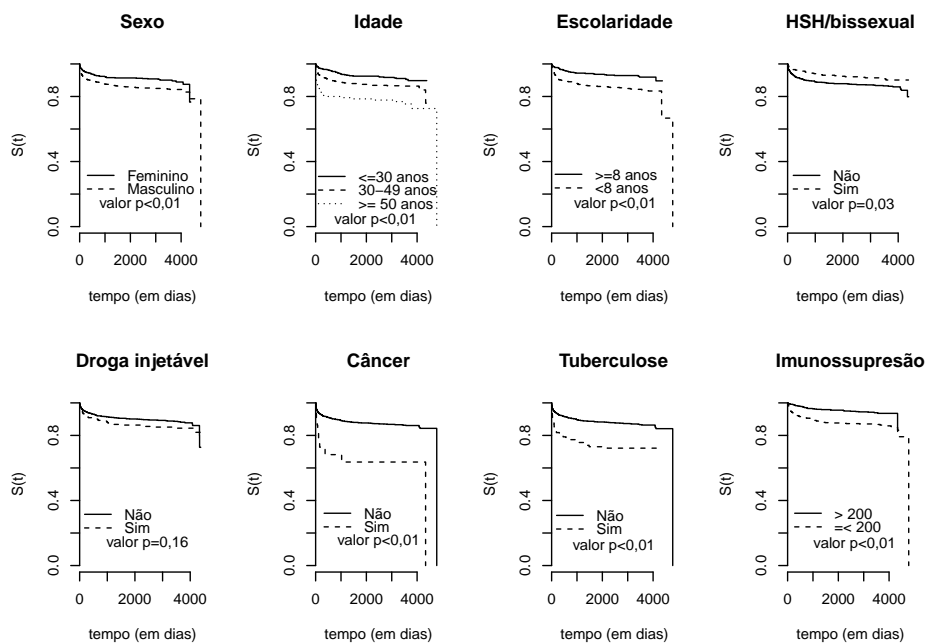


Figura 3: Curvas de sobrevivência em relação a cada variável analisada e valor p para o teste de logrank, na coorte 3 (2001-2005), Campinas, São Paulo.

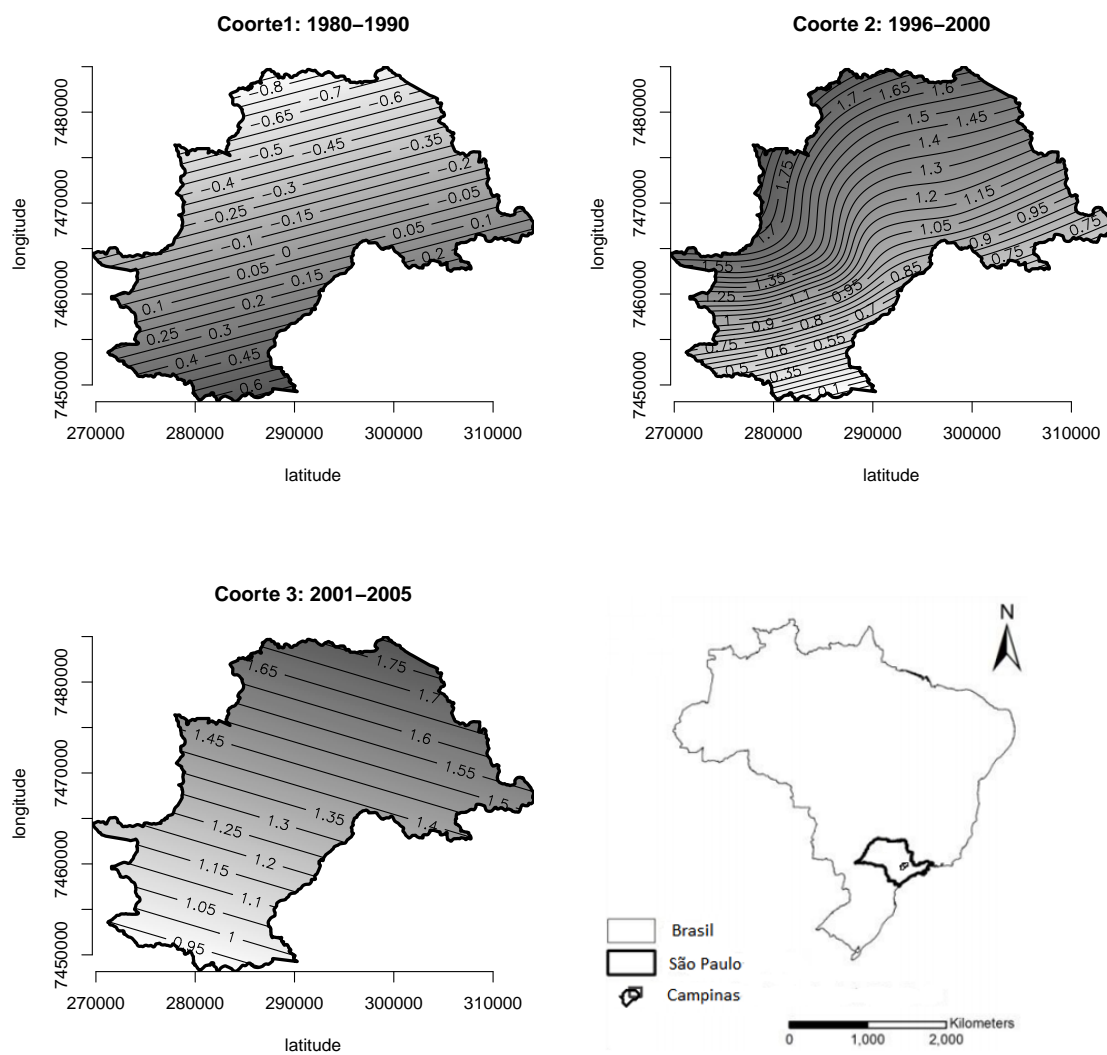


Figura 4: Mapas de risco de óbito para indivíduos diagnosticados com AIDS, obtidos do ajuste do modelo risco proporcionais aditivos, após a seleção de variáveis, nas três coortes estudadas, Campinas, São Paulo.

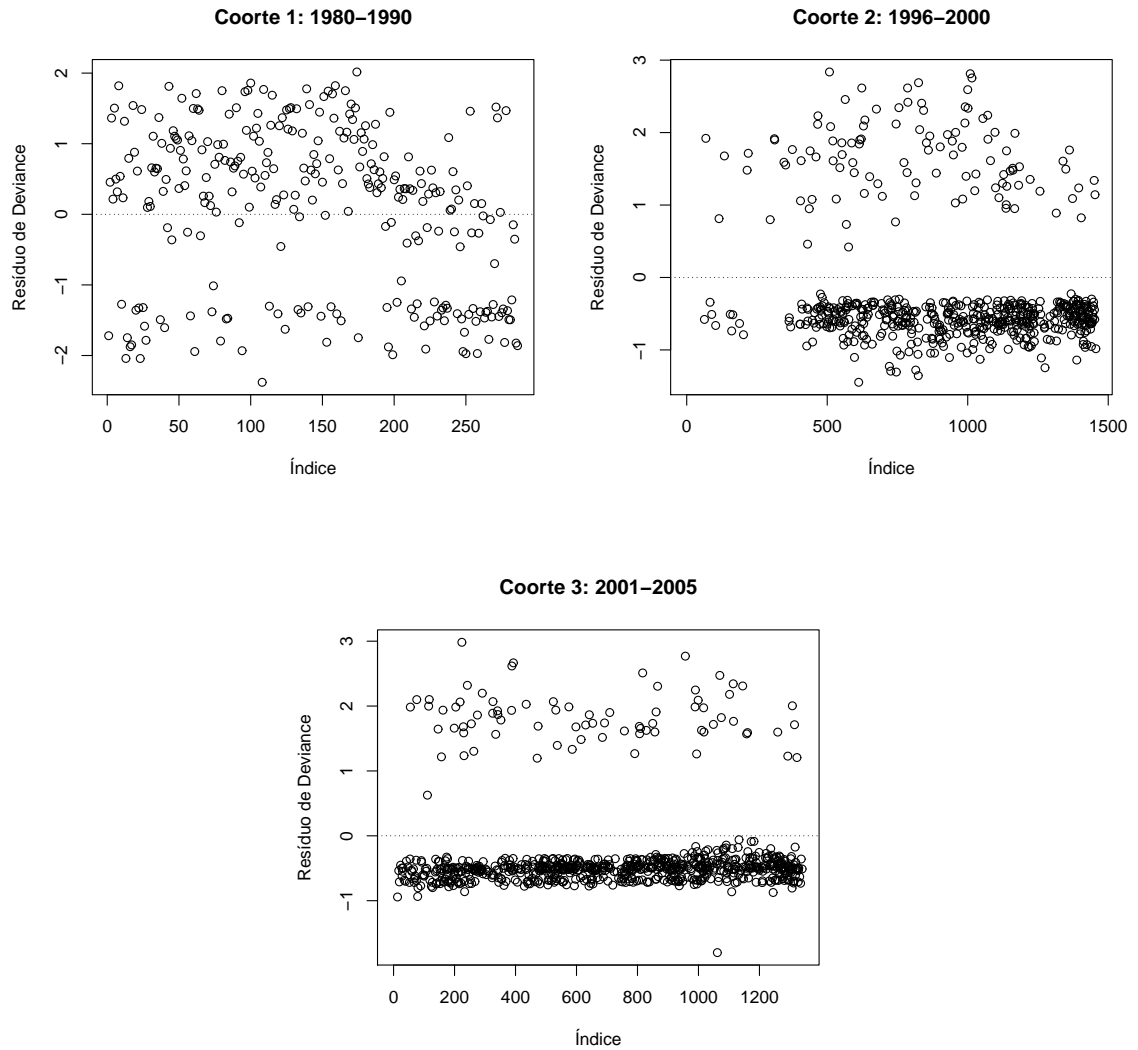


Figura 5: Resíduos de *deviance*, obtidos do ajuste dos modelos de riscos proporcionais aditivos, após a seleção de variáveis, nas três coortes estudadas em Campinas, São Paulo.

## DISCUSSÃO

O estudo das três coortes de pacientes com AIDS notificados em Campinas mostrou diferenças marcantes de sobrevivência entre a primeira década da epidemia e as coortes de indivíduos das décadas seguintes. Entretanto, observou-se variação espacial

significativa do risco de morte apenas no primeiro período analisado (1980-1990) após o ajuste do modelo multivariado, indicando a região sul de Campinas com riscos mais elevados do que as demais regiões. Apesar da variação do efeito espacial não ser significativo nas coortes 2 e coorte 3, os mapas de risco apresentam algumas áreas com estimativa pouco mais elevadas que a média do município, como na região noroeste na coorte 2 (1996-2000) e na região de fronteira ao norte da coorte 3 (2001-2005). Estas regiões apresentam níveis socioambientais precários, sendo que ao norte localiza-se o presídio da região metropolitana, onde familiares de presos residem nas redondezas. Alguns autores chamam a atenção para a maior prevalência de infecção por HIV em presidiários.<sup>21</sup> A redução do risco de morrer, particularmente após a introdução da terapia antirretroviral altamente ativa (TARV) nos serviços especializados de assistência aos pacientes com HIV-AIDS tem sido tema de vários estudos.<sup>3,12,22</sup> A distribuição 7381 casos na mesma cidade mostrou diferenças marcantes da incidência em homens e mulheres conforme as condições socioambientais dos indivíduos nas regiões do município.<sup>23</sup> Os autores sugerem diferentes perfis de infecção: ao redor do presídio, nas regiões mais ricas e na periferia da cidade, abordando períodos semelhantes aos do presente estudo. Neste, ao analisar a mortalidade, particularmente nas coortes 2 e 3, isto é, após 1996, com a disponibilização da terapia antirretroviral de alto impacto, o efeito espacial desapareceu. O acesso à medicação e ao acompanhamento pelas equipes de saúde especializadas do município possivelmente teve influência direta na sobrevivência dos pacientes. A distribuição homogênea do risco de morte nas regiões da cidade possivelmente está relacionada com a cobertura e o impacto do Programa Municipal de IST-AIDS e dos serviços de referência à AIDS no Hospital de Clínicas da UNICAMP, ambos entre os mais antigos do país, com cobertura de aproximadamente 80% dos indivíduos com HIV-AIDS moradores de Campinas. O local de moradia certamente traduz o contexto municipal socioambiental onde vive o indivíduo, entretanto nesta grande cidade do estado de São Paulo, não se mostrou associado ao risco de óbito entre os pacientes notificados, nas últimas décadas. Embora se reconheça as múltiplas dimensões das

tendências da epidemia de AIDS e as desigualdades de acesso à testagem sorológica<sup>24</sup> ao início oportuno do tratamento<sup>25</sup>, e na mortalidade em várias partes do país<sup>26</sup>, os resultados desta pesquisa sugerem que parte dos determinantes socioambientais podem ser reduzidos com a oferta e o acesso aos serviços de prevenção e assistência à AIDS.

Estudo sobre a mortalidade associada à AIDS nos distritos de São Paulo mostra que a redução ocorreu independentemente das suas condições sociais e econômicas, após a introdução da TARV no Programa municipal. Como em Campinas, observa-se a necessidade de manutenção e ampliação de políticas públicas de saúde que diminuam as desigualdades sociais ao ampliar o acesso aos cuidados e terapêuticas a parcelas amplas da população.<sup>27</sup>

Por outro lado, as curvas de Kaplan-Meier bem como o modelo aditivo multivariado de Cox mostraram diferenças significativas na sobrevida de pacientes segundo variáveis sociodemográficas e presença de comorbidades nas diferentes coortes. Na primeira década, os pacientes maiores que 30 anos apresentaram menor sobrevida, em um período com insuficientes opções terapêuticas. Nos anos 90 até 2000 do diagnóstico (coorte 2) o uso de drogas injetáveis e a “pauperização” da epidemia associaram-se à menor sobrevida dos pacientes, tendência relatada por outros autores.<sup>2,28,29</sup> Na coorte 3, última década de seguimento (2001 a 2005), observa-se maior risco de morte em homens e com baixa escolaridade. A não proporcionalidade no risco no modelo ajustado na variável imunossupressão pode indicar que, a cada ano na coorte 3, os indivíduos estão sendo diagnosticados precocemente e com uma contagem cada vez maior de células CD4. As curvas de Kaplan-Meier nesta coorte evidenciam o envelhecimento dos pacientes com HIV/AIDS, o aparecimento de comorbidades e doenças crônicas como câncer, mais frequentes que no início da epidemia, associando-se ao maior risco de complicações e à maior complexidade do quadro clínico.<sup>30</sup> Os resíduos de deviance (Figura 5), nas coortes 2 e 3, mostram uma separação clara entre os indivíduos que vieram a óbito e os indivíduos censurados. Isto pode ser explicado pelo aumento da sobrevida dos indivíduos nessas coortes, e

sugerem uma posterior análise por meio de modelos de fração de cura. A co-infecção tuberculose e AIDS tem sido relatada como associada ao abandono de tratamento de ambas as doenças. A pior evolução clínica da co-infecção pode ser devido a eventos adversos dos esquemas terapêuticos, uso de álcool, fumo, presença de doenças oportunistas e de contagem de células T-CD4 abaixo de 200 cel./mm<sup>3</sup>.<sup>31,32</sup> Algumas limitações deste estudo podem ser apontadas como a utilização de dados secundários para a obtenção de informações de incidência e mortes, nem sempre fidedignas, particularmente nas primeiras décadas da epidemia. A utilização do mapa cartográfico de 2000, pode ter dificultado a localização de endereços mais recentes, porém foi utilizado como referência intermediária do período de estudo.

## CONCLUSÃO

Observou-se marcante redução da mortalidade em pacientes com AIDS em Campinas nas várias coortes estudadas desde o início da epidemia, sugerindo impacto do acesso à medicação e aos cuidados clínicos e preventivos disponíveis no município. Por outro lado, as variáveis associadas à mortalidade no modelo multivariado foram diferentes nos distintos períodos da epidemia. Na coorte de 1980-1990 predominou as diferenças de idade, na coorte de 1996-2000 a idade mais avançada, o usuário de drogas injetáveis, tuberculose e possuir baixa contagem de células CD4 e, na coorte de 2001-2005, sexo masculino e baixa escolaridade apareceram como preditores da menor sobrevida. Além disso, o componente espacial foi significativo apenas na coorte 1, indicado que os maiores riscos de óbito ocorreram na região sul de Campinas, sendo pouco expressivo para explicar o risco de morte em pacientes com AIDS na cidade nas coortes 2 e 3, sugerindo certa homogeneidade em regiões da cidade com diferentes condições socioambientais. O estudo mostra contextos diferentes nos diversos períodos da epidemia em Campinas, indicando a necessidade de acompanhamento contínuo dos fatores de risco de morte, devido às mudanças da epidemia ao longo do tempo.

**REFERÊNCIAS**

1. UNAIDS. 90-90-90 An ambitious treatment target to help end the AIDS epidemic.; 2014. DOI: 10.7448/IAS.16.4.18751.
2. McMahon J, Wanke C, Terrin N, et al. Poverty, hunger, education, and residential status impact survival in HIV. *AIDS Behav.* 2011. DOI: 10.1007/s10461-010-9759-z.
3. Guibu IA, Barros MB de A, Donalísio MR, et al. Survival of AIDS patients in the Southeast and South of Brazil: analysis of the 1998-1999 cohort. *Cad saude publica / Minist da Saude, Fund Oswaldo Cruz, Esc Nac Saude Publica,* 2011; 27 Suppl 1: S79-S92.
4. Lopes F, Buchalla CM, Carvalho JR de, et al. Mulheres negras e não-negras e vulnerabilidade ao HIV/AIDS no estado de São Paulo, Brasil. *Rev Saúde Pública.* 2007; 41(2):39-46. DOI: 10.1590/S0034-89102007000900008.
5. Ministério da Saúde Secretaria de Vigilância em Saúde. *Boletim Epidemiológico AIDS-IST,* 2017; (5)1:8-18.
6. Grangeiro A, Castanheira ER, Nemes MIB, Grangeiro A, Castanheira ER, Nemes MIB. A re-emergência da epidemia de AIDS no Brasil: desafios e perspectivas para o seu enfrentamento. *Interface - Comun Saúde, Educ.* 2015;19(52):5-8. doi:10.1590/1807-57622015.0038.
7. Teixeira TR de A, Gracie R, Malta MS, et al. Social geography of AIDS in Brazil: identifying patterns of regional inequalities. *Cad Saude Publica.* 2014;30(2):259-271. doi:10.1590/0102-311X00051313.
8. Grangeiro A, Escuder MML, Castilho EA de. A epidemia de AIDS no Brasil e as desigualdades regionais e de oferta de serviço. *Cad Saude Publica.* 2010;26(12):2355-2367. doi:10.1590/S0102-311X2010001200014.

9. Marins JRP, Jamal LF, Chen SY, et al. Dramatic improvement in survival among adult Brazilian AIDS patients. *AIDS*. 2003; 17(11):1675-1682. DOI: 10.1097/01.AIDS.0000072649.21517.80.
10. Pereira AGL, Matos HJ de, Escosteguy CC, et al. Sobrevida de pacientes com Síndrome da Imunodeficiência Adquirida em hospital geral no Rio de Janeiro, a partir de dados da vigilância epidemiológica. *Cad Saúde Colet*. 2013; 21(2): 160-167.
11. Chequer P, Hearst N, Hudes ES, et al. Determinants of survival in adult Brazilian AIDS patients, 1982-1989. The Brazilian State AIDS Program Co-ordinators. *AIDS*, London. 1992. DOI: 10.1097/00002030-199205000-00007.
12. Lima VD, Hogg RS, Harrigan PR, et al. Continued improvement in survival among HIV-infected individuals with newer forms of highly active antiretroviral therapy. *AIDS*. 2007; 21(6): 685-692.
13. Menesia EO, Passos ADC, Monteiro ME, et al. Sobrevida de pacientes com AIDS em uma cidade do Sudeste Brasileiro. *Rev Panam Salud Pública*. 2001; 1(10): 29-36.
14. Melo MC de, Donalisio MR, Cordeiro RC. Sobrevida de pacientes com AIDS e coinfeção pelo bacilo da tuberculose nas regiões Sul e Sudeste do Brasil. *Cien Saude Colet*. 2017;22(11):3781-3792. doi:10.1590/1413-812320172211.26352015.
15. Tanaka LF, Latorre M do RDO, Gutierrez EB, et al. Cancer survival in people with AIDS: A population-based study from São Paulo, Brazil. *Int J Cancer*. 2018;142(3):524-533. doi:10.1002/ijc.31081.
16. Cox DR. Regression Models and Life-Tables. *J R Stat Soc Ser B*. 1972; 34(2): 187-220.



17. Hastie T, Tibshirani R. Generalized Additive Models. *Stat Sci.* 1986; 1(3): 297-310.
18. Hastie TJ, Tibshirani RJ. *Generalized Additive Models*. 1st ed. Boca Raton, Florida: Chapman & Hall/CRC, 1990.
19. Wood SN. *Generalized Additive Models: An Introduction with R*. Boca Raton, Florida: Chapman & Hall/CRC, 2006.
20. R Core Team. R: A Language and Environment for Statistical Computing <https://www.r-project.org/> (2017).
21. de Albuquerque ACC, da Silva DM, Cabral Rabelo DC, et al. Soroprevalência e fatores associados ao Vírus da Imunodeficiência Humana (HIV) e sífilis em presidiários do Estado de Pernambuco, Brasil. *Rev Ciência Saúde Coletiva*. 2014; 19(7):2125:2132.
22. Deeks SG, Lewin SR, Havlir D V. The end of AIDS: HIV infection as a chronic disease. *Lancet*. 2013; 382(9903): 1525-1533.
23. Stephan C, Henn CA, Donalisio MR. Expressão geográfica da Epidemia de AIDS em Campinas, São Paulo, de 1980 a 2005. *Rev Saúde Pública*. 2010; 44(5): 812-819.
24. França Junior I, Calazans G, Zucchi EM. Mudanças no âmbito da testagem anti-HIV no Brasil entre 1998 e 2005. *Rev Saúde Pública*. 2008;42(suppl 1):84-97. doi:10.1590/S0034-89102008000800011.
25. Souza-Jr PRB, Szwarcwald CL, Castilho EA. Delay in introducing antiretroviral therapy in patients infected by HIV in Brazil, 2003-2006. *Clinics*. 2007;62(5):579-584. doi:10.1590/S1807-59322007000500008.
26. Dourado I, Veras MA de SM, Barreira D, Brito AM de. Tendências da epidemia de AIDS no Brasil após a terapia anti-retroviral. *Rev Saúde Pública*. 2006;40(suppl):9-17. doi:10.1590/S0034-89102006000800003.

27. Antunes JLF, Waldman EA, Borrell C. Is it possible to reduce AIDS deaths without reinforcing socioeconomic inequalities in health? *In J Epidemiol.* 2005; 34(3): 586-592.
28. Santos NJS, Tayra A, Silva SR, et al. A AIDS no Estado de São Paulo: as mudanças no perfil da epidemia e perspectivas da vigilância epidemiológica. *Rev Bras Epidemiol.* 2002; 5(3): 286-310.
29. Farias NE, Chester LGC. Tendências da morbimortalidade por AIDS e condições socioeconômicas no município de São Paulo, 1994 a 2001. *Rev Bras Epidemiol.* 2004; 7(4): 489-502.
30. Schneider MF, Gange SJ, Williams CM, et al. Patterns of the hazard of death after AIDS through the evolution of antiretroviral therapy: 1984-2004. *AIDS.* 2005; 19(17): 489-502. DOI: 10.1097/01.AIDS.0000189864.90053.22.
31. Maruza M, Albuquerque MFPM, Coimbra I, et al. Risk factors for default from tuberculosis treatment in HIV-infected individuals in the state of Pernambuco, Brazil: a prospective cohort study. *BMC Infect Dis.* 2011; 11(1): 351.
32. Whalen CC, Nsubuga P, Okwera A, et al. Impact of pulmonary tuberculosis on survival of HIV-infected adults: a prospective epidemiologic study in Uganda. *AIDS.* 2000; 14(9): 1219-1228.

## 4 DISCUSSÃO E CONCLUSÃO GERAL

A utilização de técnicas de análise de sobrevivência e de estatística espacial em pesquisas na área de saúde vem crescendo nas últimas décadas, com o objetivo de descrever os fatores de riscos que influenciam na sobrevivência dos indivíduos, na identificação de padrões espaciais e na identificação de grupos de maior risco a uma determinada doença. Apesar dessas técnicas serem importantes no monitoramento da doença e auxiliar em ações na tomada de decisão do gestor de saúde, o uso delas concomitantemente em epidemiologia é ainda incipiente na obtenção de indicadores de risco.

A partir dessas considerações, esta tese buscou reunir estas duas ferramentas, estatística espacial e análise de sobrevivência, com o intuito de avaliar a sobrevivência e o risco espacial de óbito de pacientes portadores de HIV/AIDS residentes no município de Campinas, em três períodos de tempo, por meio do ajuste de modelos de sobrevivência espaço-temporal, obtendo mapas que servem como indicadores de risco de óbito.

No primeiro artigo, realizou-se o ajuste do modelo semi-paramétrico bayesiano empregando o método de aproximação de Laplace Aninhada e Integrada (INLA), segundo os setores censitários. Os resultados evidenciaram maior risco de óbito em homens, idosos e na categoria transmissão heterossexual. Além disso, os mapas obtidos indicaram que as regiões de maior carência e piores condições socioambientais coincidiam com as regiões de maior risco de morrer na cidade de Campinas, principalmente nas coortes 1 e 2.

No segundo artigo, buscou-se trabalhar uma forma alternativa aos modelos bayesianos. A partir deste propósito empregou-se uma extensão do modelo de

riscos proporcionais, incorporado nos modelos aditivos generalizados (GAM), considerando uma função de splines para modelar as localizações espaciais dos indivíduos (latitude e longitude). Neste artigo foi possível atualizar o tempo de seguimento na última coorte (2001-2005) de 2012 para o ano de 2014, em decorrência do acesso a base de dados desse período no Sistema de Informação sobre Mortalidade (SIM) de Campinas, e também, acrescentar na análise duas outras variáveis como tuberculose e câncer. Como resultados, destaca-se a importância da TARV na sobrevivência dos indivíduos ao longo do tempo. Na última coorte analisada, observam-se como fatores de risco ser homem e ter baixa escolaridade, sendo este modelo estratificado pela variável imunossupressão grave. No modelo GAM, empregado, o componente espacial foi estatisticamente significativo ao nível de 5% de significância, apenas na coorte 1, sugerindo na coorte 2 e 3 seguimento clínico e acesso à TARV com certa homogeneidade em regiões da cidade com diferentes condições socioambientais.

Portanto, os resultados apresentados nestes dois artigos evidenciaram: a relevância de se estudar a epidemia da AIDS periodicamente, afim de entender a dinâmica da epidemia como evidenciou os períodos analisados que resultaram em conjunturas epidemiológicas distintas. A necessidade de aplicar modelos adequados, que permitem analisar o risco em função do tempo até o óbito de um indivíduo condicionado às covariáveis do problema e do sistema de informação geográfica. E o uso do método estimação INLA em modelos hierárquicos bayesianos e dos modelos aditivos generalizados (GAM), que ainda são poucos explorados em trabalhos de epidemiologia.

A partir da análise de resíduos do modelo GAM, observa-se nas duas últimas coortes uma separação dos indivíduos que falharam dos indivíduos censurados, o que pode ser explicado pela queda drástica da mortalidade nessas duas coortes, sugerindo tratar em trabalhos futuros esses dados por meio de modelos de fração de cura.

## REFERÊNCIAS BIBLIOGRÁFICAS

BESAG, J.; KOOPERBERG, C. On conditional and intrinsic autoregressions. **Biometrika**, v.82, n.4, p.733–746, 1995.

BLANGIARDO, M.; CAMELETTI, M. **Spatial and Spatio-temporal Bayesian Models with R - INLA**. West Sussex: Wiley, 2015. 308p.

BOLETIM EPIDEMIOLÓGICO CRT- PE – DST/AIDS/CVE. Secretária de estado da Saúde de São Paulo, São Paulo, Brasil, 2015. ISBN 1984-64x.

BOLETIM EPIDEMIOLÓGICO HIV/AIDS. Secretária de Vigilância e Saúde – Ministério da Saúde, Brasília, Brasil, 2016. ISBN 2358-9450.

BRESLOW, N. Contribution to the discussion of the paper by D.R Cox. **Journal of the Royal Statistical Society B**, v.34, n.1, p.216–217, 1972.

CARLIN, B. P.; LOUIS, T. A. **Bayesian Methods for Analysis**. 3. ed. New York: Chapman & Hall/CRC, 2009. 520p.

CARVALHO, M.; SOUZA-SANTOS, R. Análise de dados espaciais em saúde pública: métodos, problemas, perspectivas. **Cad. Saúde Pública**, v.21, n.2, p.361–378, 2005.

CHEQUER, P.; HEARST, N.; HUDES, E. S.; CASTILHO, E.; RUTHERFORD, G.; LOURES, L.; RODRIGUES, L. Determinants of survival in adult Brazilian AIDS patients, 1982-1989. The Brazilian State AIDS Program Co-Ordinators. **AIDS (London, England)**, v.6, n.5, p.483–487, 1992.

COLOSIMO, E. A.; GIOLO, S. R. **Análise de sobrevivência aplicada**. São Paulo: Edgard Blücher, 2006. 392p.

COX, D. R. Regression models and life-tables. **Journal of the Royal statistical Society-B**, v.34, n.1, p.187–220, 1972.

COX, D. R. Partial likelihood. **Biometrika**, v.62, n.2, p.269–276, 1975.

FONSECA, M. G. P.; BASTOS, F. I. Twenty-five years of the AIDS epidemic in Brazil: principal epidemiological findings, 1980-2005. **Cadernos de saude publica / Ministerio da Saude, Fundacao Oswaldo Cruz, Escola Nacional de Saude Publica**, 2007.

GOTTLIEB, M. S.; SCHROFF, R.; SCHANKER, H. M.; WEISMAN, J. D.; FAN, P. T.; WOLF, R. A.; SAXON, A. Pneumocystis carinii pneumonia and mucosal candidiasis in previously healthy homosexual men: evidence of a new acquired cellular immunodeficiency. **New England Journal of Medicine**, v.305, n.24, p.1425–1431, 1981.

GUIBU, I. A.; BARROS, M. B. D. A.; DONALÍSIO, M. R.; TAYRA, Â.; ALVES, M. C. G. P. Survival of AIDS patients in the Southeast and South of Brazil: analysis of the 1998-1999 cohort. **Cadernos de saude publica / Ministerio da Saude, Fundacao Oswaldo Cruz, Escola Nacional de Saude Publica**, v.27 Suppl 1, p.S79–S92, 2011.

HASTIE, T.; TIBSHIRANI, R. Generalized Additive Models. **Statistical Science**, v.1, n.3, p.297–310, 1986.

HASTIE, T.; TIBSHIRANI, R. **Generalized Additive Models**. Boca Raton, Florida: Chapman & Hall/CRC, 1990. 329p.

KILLIAN, M. S.; LEVY, J. A. HIV/AIDS: 30 years of progress and future challenges. **European journal of immunology**, v.41, n.12, p.3401–3411, 2011.

LAWLESS, J. F. **Statistical models and methods for lifetime data**. New Jersey: Wiley, 2002. 621p.

LIMA, V. D.; HOGG, R. S.; HARRIGAN, P. R.; MOORE, D.; YIP, B.; WOOD, E.; MONTANER, J. S. G. Continued improvement in survival among HIV-infected individuals with newer forms of highly active antiretroviral therapy. **AIDS (London, England)**, v.21, n.6, p.685–692, 2007.

MARTINO, S.; AKERKAR, R. Approximate Bayesian Inference for Survival Models. **Statistics**, 2010.

MCMAHON, J.; WANKE, C.; TERRIN, N.; SKINNER, S.; KNOX, T. Poverty, hunger, education, and residential status impact survival in HIV. **AIDS and Behavior**, 2011.

MENESIA, E. O.; PASSOS, A. D. C.; MONTEIRO, M. E.; DAL-FABBRO, A. L.; LAPREGA, M. R. Sobrevivência de pacientes com AIDS em uma cidade do Sudeste Brasileiro. **Rev. Panam. Salud Pública**, v.1, n.10, p.29–36, 2001.

MINISTÉRIO DA SAÚDE. O que é o HIV? <http://www.aids.gov.br/pt-br/publico-geral/o-que-e-hiv>, 2017. [Online; acessado 16-AGOSTO-2017].

NELDER, J.; WEDDERBURN, R. General linearized models. **J. Roy. Stat. Soc. Ser. A**, v.135, p.370–384, 1972.

PETO, R. Contribution to the discussion of the paper by D.R Cox. **Journal of the Royal Statistical Society B**, v.34, n.1, p.205–207, 1972.

R DEVELOPMENT CORE TEAM. **R: A Language and Environment for Statistical Computing**. R Foundation for Statistical Computing, Vienna, Austria, 2016. ISBN 3-900051-08-9.

RIBEIRO JR, P. J.; BONAT, W. H.; KRAINSKI, E. T.; ZEVIANI, W. M. Métodos Computacionais em Inferência Estatística. In: 20º SINAPE - SIMPÓSIO NACIONAL DE PROBABILIDADE E ESTATÍSTICA, João Pessoa, 2012. **Minicurso**; resumos. João Pessoa: ABE, 2012. 277.

RUE, H.; MARTINO, S.; CHOPIN, N. Approximate Bayesian inference for latent Gaussian models by using integrated nested Laplace approximations. **Journal of the Royal Statistical Society. Series B: Statistical Methodology**, v.71, n.2, p.319–392, 2009.

SANTOS, N. J. S.; TAYRA, A.; SILVA, S. R.; BUCHALLA, C. M.; LAURENTI, R. A aids no Estado de São Paulo: as mudanças no perfil da epidemia e perspectivas da vigilância epidemiológica. **Revista Brasileira de Epidemiologia**, v.5, n.3, p.286–310, 2002.

SCHOENFELD, D. Partial Residuals for The Proportional Hazards Regression Model. **Biometrika**, v.69, n.1, p.239–241, 1982.

SOUSA, A. I. A. D.; PINTO JÚNIOR, V. L. Análise espacial e temporal dos casos de aids no Brasil em 1996-2011: áreas de risco aumentado ao longo do tempo. **Epidemiologia e Serviços de Saúde**, v.25, n.3, p.467–476, 2016.

UNAIDS. 90-90-90 An ambitious treatment target to help end the AIDS epidemic. Rel. téc., 2014.

WOOD, S. N. **Generalized Additive Models: An Introduction with R**. Boca Raton, Florida: Chapman & Hall/CRC, 2006. 391p.



# APÊNDICES

## Apêndice A

### MODELO SOBREVIVÊNCIA ESPAÇO – TEMPORAL USANDO INLA

```
setwd("E:/MATERIAS_PROJETO E TESE DOUT_BIOMETIRA/ANALISES_DOUT_2016")  
## PACOTES ##  
#install.packages("INLA", repos=c(getOption("repos"),  
INLA="https://inla.r-inla-download.org/R/stable"), dep=TRUE)  
#install.packages("INLA", repos=c(getOption("repos"),  
INLA="https://inla.r-inla-download.org/R/testing"), dep=TRUE)  
require(maptools)  
require(spdep)  
require(lattice)  
require(rgdal)  
require(INLA)  
##### COMANDO PARA ENTRAR COM SHAPELFILE ####  
aids=readOGR("mapa_camp_agregado.shp")
```

CRIANDO UM ARQUIVO DA MATRIZ DE VIZINHANÇA PARA USO NO MO-  
DELO LATENTE

```

temp <- poly2nb(aids,queen=T)
nb2INLA("LDN.graph", temp)
## SALVANDO O ARQUIVO NA PASTA ##
LDN.adj <- paste(getwd(),"/LDN.graph",sep="")
## ANALISE COORTE 1 ##
dados<-read.csv2("coorte1.csv",h=T,dec=",")
ordenado=order(dados$ID_LISA)
ordenado
dados=dados[ordenado,]
dados$idade<-ifelse(dados$idade_diag<=30,0,ifelse(dados$idade_diag>30 &
dados$idade_diag<50,1,2))
## FATORES ##
dados$idade2<-factor(dados$idade)
dados$sexo2<-factor(dados$sexo)

AJUSTANDO O MODELO DE SOBREVIVÊNCIA

formula = inla.surv(dados$tempo,dados$censura) ~ sexo2+idade2+escol+
droga_inj+hom_het+f(ID_LISA,model="besag",graph=LDN.adj)
modelo= inla(formula, family="coxph", data=dados,verbose=TRUE,
control.hazard=list(model="rw1",n.intervals=20,
param=c(1,0.001)),control.predictor=list(compute=TRUE))
summary(modelo)
round(modelo$summary.fixed,3)

GRÁFICOS

X11()
plot(modelo)
## MAPA DE RISCO PARA CADA UM DOS SETORES CENSITARIOS ##
aids$efeitoarea1<-modelo$summary.random$ID_LISA$mean
x11()

```

```

zeta.ST1 <- unlist(aids$efeitoarea1)
quantile(zeta.ST1)
SMR.cutoff<- c(quantile(zeta.ST1))
xi.factor <- cut(zeta.ST1,breaks=SMR.cutoff,include.lowest=TRUE)
data.aids <- attr(aids, "data")
attr(aids, "data")=data.frame(data.aids, xi=xi.factor)
trellis.par.set(axis.line=list(col=NA))
spplot(obj=aids, zcol= "xi", col.regions=gray(3.5:0.5/4),
main="Coorte 1: 1980-1990")

## SEM COVARIAVEL - coorte 1 - 1980-1990 ##
formula = inla.surv(dados$tempo,dados$censura) ~
f(ID_LISA,model="besag",graph=LDN.adj)
modelo= inla(formula, family="coxph", data=dados,verbose=TRUE,
              control.hazard=list(model="rw1",n.intervals=20,
param=c(1,0.001)),control.predictor=list(compute=TRUE))
summary(modelo)
round(modelo$summary.fixed,3)

```

Obs.: As rotinas para as coortes 2 e 3 são as mesmas apenas sendo necessário modificar os bancos de dados:

```

## COORTE 2 - 1996-2000 ##
dados<-read.csv2("coorte2.csv",h=T,dec=",")
## COORTE 3 - 2001-2005 ##
dados<-read.csv2("coorte3.csv",h=T,dec=",")

```

## Apêndice B

### MODELO SOBREVIVÊNCIA ESPAÇO – TEMPORAL VIA GAM

```

# PACOTES #
#install.packages("sandwich")
#install.packages(c("maptools","mgcv","survival","rgdal",
#"ggplot2","Rcmdr","splancs"))
sapply(c("maptools","mgcv","survival","Rcmdr",
"rgdal","ggplot2","splancs"),require,char=T)
# DADOS #
geral <-readXL("I:/MATERIAS_PROJETO E TESE DOUT_BIOMETIRA/
ANALISES_DOUT_2016/ANALISE_COX/GERAL_MOD2.xlsx",rownames=FALSE,
header=TRUE, na="", sheet="GERAL_MOD2", stringsAsFactors=TRUE)
geral <- within(geral, {
droga_inj<-as.factor(droga_inj)
escol<-as.factor(escol)
sexo<-as.factor(sexo)
hom_het<-as.factor(hom_het)
ANT_TUBERC <- as.factor(ANT_TUBERC)
idade1<- as.factor(idade1)
imupress<-as.factor(imupress)
cancer<-as.factor(cancer)
})
dados1 <- subset(geral, subset=coorte=="coorte1")
dados2<-subset(geral,subset=coorte=="coorte3")
dados3<-subset(geral,subset=coorte=="coorte4")
## MAPA DE CAMPINAS ##
camp.map <- readOGR(dsn="I:/MATERIAS_PROJETO E TESE DOUT_BIOMETIRA/
ANALISES_DOUT_2016/ANALISE_COX", layer="camp21")

```

```

## CONVERTER O SHAPEFILE EM PONTOS ##
camp.map_df <- fortify(camp.map)
coord_camp<-camp.map_df[,c(1,2)]
coord_camp2<-coord_camp[c(1:3356),]
plot(coord_camp2)
## CURVAS DE SOBREVIVENCIA PARA COORTE 1 ##
par(mfrow=c(2,4))
kmsexo<-survfit(Surv(tempo,censura)~sexo,data=dados1)
plot(kmsexo,col=1,lty=1:2,mark.time=F,main="Sexo",bty="n",
xlab="tempo (em dias)",ylab="S(t)")
legend(locator(1),legend=c("Feminino","Masculino"),
col=1,lty=c(1,2),bty="n",cex=1)
legend(locator(1),legend=c("valor p=0,79"),col=1,lty=0,
cex=1,bty="n")
kmidade<-survfit(Surv(tempo,censura)~idade1,data=dados1)
plot(kmidade,lty=1:3,col=1,bty="n",xlab="tempo (em dias)",
ylab="S(t)",main="Idade")
legend(locator(1),col=1,lty=1:3,legend=c("<=30 anos",
"30-49 anos",">= 50 anos"),bty="n",cex=1)
legend(locator(1),legend=c("valor p<0,01"),col=1,lty=0,
cex=1,bty="n")
kmesc<-survfit(Surv(tempo,censura)~escol,data=dados1)
plot(kmesc,lty=1:2,col=1,bty="n",xlab="tempo (em dias)",
ylab="S(t)",main="Escolaridade")
legend(locator(1),lty=1:2,col=1,legend=c(">=8 anos","<8 anos"),
bty="n",cex=1)
legend(locator(1),legend=c("valor p=0,48"),col=1,lty=0,
cex=1,bty="n")
kmhom<-survfit(Surv(tempo,censura)~hom_het,data=dados1)

```

```

plot(kmhom,lty=1:2,col=1,bty="n",xlab="tempo (em dias)",
ylab="S(t)",main="HSH/bissexual")
legend(locator(1),lty=1:2,col=1,legend=c("Nao","Sim"),bty="n",cex=1)
legend(locator(1),legend=c("valor p=0,13"),col=1,lty=0,cex=1,bty="n")
kmdrog<-survfit(Surv(tempo,censura)~droga_inj,data=dados1)
plot(kmdrog,lty=1:2,col=1,bty="n",xlab="tempo (em dias)",
ylab="S(t)",main="Droga injetavel")
legend(locator(1),lty=1:2,col=1,legend=c("Nao","Sim"),bty="n",cex=1)
legend(locator(1),legend=c("valor p=0,56"),col=1,lty=0,cex=1,bty="n")
kmcancer<-survfit(Surv(tempo,censura)~cancer,data=dados1)
plot(kmcancer,lty=1:2,col=1,bty="n",xlab="tempo (em dias)",
ylab="S(t)",main="Cancer")
legend(locator(1),lty=1:2,col=1,legend=c("Nao","Sim"),bty="n",cex=1)
legend(locator(1),legend=c("valor p=0,52"),col=1,lty=0,cex=1,bty="n")
kmtuber<-survfit(Surv(tempo,censura)~ANT_TUBERC,data=dados1)
plot(kmtuber,lty=1:2,col=1,bty="n",xlab="tempo (em dias)",
ylab="S(t)",main="Tuberculose")
legend(locator(1),lty=1:2,col=1,legend=c("Nao","Sim"),bty="n",cex=1)
legend(locator(1),legend=c("valor p=0,18"),col=1,lty=0,cex=1,bty="n")
kmimun<-survfit(Surv(tempo,censura)~imupress,data=dados1)
plot(kmimun,lty=1:2,col=1,bty="n",xlab="tempo (em dias)",
ylab="S(t)",main="Imunossupresao")
legend(locator(1),lty=1:2,col=1,legend=c("> 200","=< 200"),bty="n",cex=1)
legend(locator(1),legend=c("valor p=0,18"),col=1,lty=0,cex=1,bty="n")
## AJUSTE DO MODELO DE COX COMPLETO COORTE 1 ##
mod1<-coxph(Surv(tempo1,censura)~sexo+escol+hom_het+
idade1+droga_inj+cancer+ANT_TUBERC,data=dados1)
summary(mod1)
## TESTANDO A SUPOSICAO DE RISCOS PROPORCIONAIS ##

```

```

cox.zph(mod1)
## AJUSTE DO MODELO GAM ##
## MODELO PARA COORTE 1 ##
gamtp1<-gam(tempo~sexo+escol+hom_het+idade1+droga_inj+cancer+ANT_TUBERC+
s(POINT_X,POINT_Y),family=cox.ph(),weights=censura,data=dados1)
### MODELO SELECIONADO ###
gamtp1<-gam(tempo~idade1+s(POINT_X,POINT_Y),
family=cox.ph(),weights=censura,data=dados1)
summary(gamtp1)
## AVALIANDO OS RESIDUOS MARTINGALE E DEVIANCE ####
resdev<-resid(gamtp1,type="deviance")
plot(resdev,xlab="Indice",ylab="Residuo de Deviance",
main="Coorte 1: 1980-1990")
abline(h=0,lty=3)
## MAPA DE RISCO ##
## PROGRAMA AUXILIAR PARA LIMITAR OS RISCO NOS LIMITES DO SHAPEFILE ##
vis.gam<-function (x, view = NULL, cond = list(),
n.grid = 30, too.far = 0,
col = NA, color = "heat", contour.col = NULL, se = -1, type = "link",
plot.type = "persp", zlim = NULL, nCol = 50, longlim, latlim,...)
{
fac.seq <- function(fac, n.grid) {
fn <- length(levels(fac))
gn <- n.grid
if (fn > gn)
mf <- factor(levels(fac))[1:gn]
else {
ln <- floor(gn/fn)
mf <- rep(levels(fac)[fn], gn)

```

```
mf[1:(ln * fn)] <- rep(levels(fac), rep(ln, fn))
mf <- factor(mf, levels = levels(fac))
}
mf
}
dnm <- names(list(...))
v.names <- names(x$var.summary)
if (is.null(view)) {
k <- 0
view <- rep("", 2)
for (i in 1:length(v.names)) {
ok <- TRUE
if (is.matrix(x$var.summary[[i]]))
ok <- FALSE
else if (is.factor(x$var.summary[[i]])) {
if (length(levels(x$var.summary[[i]])) <= 1)
ok <- FALSE
}
else {
if (length(unique(x$var.summary[[i]])) == 1)
ok <- FALSE
}
if (ok) {
k <- k + 1
view[k] <- v.names[i]
}
if (k == 2)
break
}
}
```



```

if (k < 2)
stop("Model does not seem to have enough terms to do
anything useful")
  }
else {
if (sum(view %in% v.names) != 2)
stop(gettextf("view variables must be one of %s",
paste(v.names, collapse = ", ")))
for (i in 1:2) if (!inherits(x$var.summary[[view[i]]],
c("numeric", "factor")))
stop("Don't know what to do with parametric terms that
are not simple numeric or factor variables")
  }
ok <- TRUE
for (i in 1:2) if (is.factor(x$var.summary[[view[i]]])) {
if (length(levels(x$var.summary[[view[i]]])) <= 1)
ok <- FALSE
  }
else {
if (length(unique(x$var.summary[[view[i]]])) <= 1)
ok <- FALSE
  }
if (!ok)
stop(gettextf("View variables must contain more than one
value. view = c(%s,%s).",
view[1], view[2]))
if (is.factor(x$var.summary[[view[1]]]))
m1 <- fac.seq(x$var.summary[[view[1]]], n.grid)
else {

```

```

r1 <- longlim
m1 <- seq(r1[1], r1[2], length = n.grid)
  }
if (is.factor(x$var.summary[[view[2]]]))
m2 <- fac.seq(x$var.summary[[view[2]]], n.grid)
else {
r2 <- latlim
m2 <- seq(r2[1], r2[2], length = n.grid)
  }
v1 <- rep(m1, n.grid)
v2 <- rep(m2, rep(n.grid, n.grid))
newd <- data.frame(matrix(0, n.grid * n.grid, 0))
for (i in 1:length(x$var.summary)) {
ma <- cond[[v.names[i]]]
if (is.null(ma)) {
ma <- x$var.summary[[i]]
if (is.numeric(ma))
ma <- ma[2]
  }
if (is.matrix(x$var.summary[[i]]))
newd[[i]] <- matrix(ma, n.grid * n.grid,
ncol(x$var.summary[[i]]), byrow = TRUE)
else newd[[i]] <- rep(ma, n.grid * n.grid)
  }
names(newd) <- v.names
newd[[view[1]]] <- v1
newd[[view[2]]] <- v2
if (type == "link")
zlab <- paste("linear predictor")

```

```

else if (type == "response")
zlab <- type
else stop("type must be \"link\" or \"response\"")
fv <- predict.gam(x, newdata = newd, se.fit = TRUE, type = type)
z <- fv$fit
if (too.far > 0) {
ex.tf <- exclude.too.far(v1, v2, x$model[, view[1]],
x$model[, view[2]], dist = too.far)
fv$se.fit[ex.tf] <- fv$fit[ex.tf] <- NA
}
if (is.factor(m1)) {
m1 <- as.numeric(m1)
m1 <- seq(min(m1) - 0.5, max(m1) + 0.5, length = n.grid)
}
if (is.factor(m2)) {
m2 <- as.numeric(m2)
m2 <- seq(min(m1) - 0.5, max(m2) + 0.5, length = n.grid)
}
if (se <= 0) {
old.warn <- options(warn = -1)
av <- matrix(c(0.5, 0.5, rep(0, n.grid - 1)), n.grid,
n.grid - 1)
options(old.warn)
max.z <- max(z, na.rm = TRUE)
z[is.na(z)] <- max.z * 10000
z <- matrix(z, n.grid, n.grid)
surf.col <- t(av) %*% z %*% av
surf.col[surf.col > max.z * 2] <- NA
if (!is.null(zlim)) {

```

```
if (length(zlim) != 2 || zlim[1] >= zlim[2])
  stop("Something wrong with zlim")
min.z <- zlim[1]
max.z <- zlim[2]
}
else {
  min.z <- min(fv$fit, na.rm = TRUE)
  max.z <- max(fv$fit, na.rm = TRUE)
}
surf.col <- surf.col - min.z
surf.col <- surf.col/(max.z - min.z)
surf.col <- round(surf.col * nCol)
con.col <- 1
if (color == "heat") {
  pal <- heat.colors(nCol)
  con.col <- 3
}
else if (color == "topo") {
  pal <- topo.colors(nCol)
  con.col <- 2
}
else if (color == "cm") {
  pal <- cm.colors(nCol)
  con.col <- 1
}
else if (color == "terrain") {
  pal <- terrain.colors(nCol)
  con.col <- 2
}
```

```

else if (color == "gray" || color == "bw") {
  pal <- gray(seq(0.1, 0.9, length = nCol))
  con.col <- 1
}
else stop("color scheme not recognised")
if (is.null(contour.col))
  contour.col <- con.col
surf.col[surf.col < 1] <- 1
surf.col[surf.col > nCol] <- nCol
if (is.na(col))
  col <- pal[as.array(surf.col)]
z <- matrix(fv$fit, n.grid, n.grid)
if (plot.type == "contour") {
  stub <- paste(iffelse("xlab" %in% dnm, "", ",xlab=view[1]"),
    iffelse("ylab" %in% dnm, "", ",ylab=view[2]"),
    iffelse("main" %in% dnm, "", ",main=zlab"), ",...)",
    sep = "")
  if (color != "bw") {
    txt <- paste("image(m1,m2,z,col=pal,
zlim=c(min.z,max.z)",
      stub, sep = "")
    eval(parse(text = txt))
  }
  print(paste(1,txt))

  txt <- paste("contour(m1,m2,z,
col=contour.col,zlim=c(min.z,max.z)",
    iffelse("add" %in% dnm, "", ",add=TRUE"), ",...)",
    sep = "")
  eval(parse(text = txt))
}

```

```

print(paste(2,txt))
  }
  else {
      txt <- paste("contour(m1,m2,z,col=1,
zlim=c(min.z,max.z)",
      stub, sep = "")
      eval(parse(text = txt))
print(paste(3,txt))
  }
}
else {
      stub <- paste(iffelse("xlab" %in% dnm, "", ",xlab=view[1]"),
      iffelse("ylab" %in% dnm, "", ",ylab=view[2]"),
      iffelse("zlab" %in% dnm, "", ",zlab=zlab"), ",...)",
      sep = "")
      if (color == "bw") {
          op <- par(bg = "white")
          txt <- paste("persp(m1,m2,z,
col=\"white\",zlim=c(min.z,max.z) ",
          stub, sep = "")
          eval(parse(text = txt))
print(paste(4,txt))
      par(op)
  }
  else {
      txt <- paste("persp(m1,m2,z,col=col,
zlim=c(min.z,max.z)",
      stub, sep = "")
      eval(parse(text = txt))

```

```

print(paste(5,txt))
    }
  }
}
else {
  if (color == "bw" || color == "gray") {
    subs <- paste("grey are +/-", se, "s.e.")
    lo.col <- "gray"
    hi.col <- "gray"
  }
  else {
    subs <- paste("red/green are +/-", se, "s.e.")
    lo.col <- "green"
    hi.col <- "red"
  }
  if (!is.null(zlim)) {
    if (length(zlim) != 2 || zlim[1] >= zlim[2])
      stop("Something wrong with zlim")
    min.z <- zlim[1]
    max.z <- zlim[2]
  }
  else {
    max.z <- max(fv$fit + fv$se.fit * se, na.rm = TRUE)
    min.z <- min(fv$fit - fv$se.fit * se, na.rm = TRUE)
    zlim <- c(min.z, max.z)
  }
  z <- fv$fit - fv$se.fit * se
  z <- matrix(z, n.grid, n.grid)
  if (plot.type == "contour")

```

```

warning("sorry no option for
contouring with errors: try plot.gam")
stub <- paste(iffelse("xlab" %in% dnm, "", ",xlab=view[1]"),
iffelse("ylab" %in% dnm, "", ",ylab=view[2]"), iffelse("zlab" %in%
dnm, "", ",zlab=zlab"), iffelse("sub" %in% dnm,
"", ",sub=subs"), ",...)", sep = "")
txt <- paste("persp(m1,m2,z,col=col,zlim=zlim", iffelse("border" %in%
dnm, "", ",border=lo.col"), stub, sep = "")
eval(parse(text = txt))
print(paste(6,txt))
par(new = TRUE)
z <- fv$fit
z <- matrix(z, n.grid, n.grid)
txt <- paste("persp(m1,m2,z,col=col,zlim=zlim",
iffelse("border" %in%
dnm, "", ",border=\"black\""), stub, sep = "")
eval(parse(text = txt))
print(paste(7,txt))
par(new = TRUE)
z <- fv$fit + se * fv$se.fit
z <- matrix(z, n.grid, n.grid)
txt <- paste("persp(m1,m2,z,col=col,zlim=zlim",
iffelse("border" %in%
dnm, "", ",border=hi.col"), stub, sep = "")
eval(parse(text = txt))
print(paste(8,txt))
}
return(list(m1,m2,z))
}

```



PARA A NOVA FUNÇÃO `vis.gam`, ROTINA ACIMA, CRIAMOS UM ARQUIVO SEPARADO SENDO CARREGADO PELA SEGUINTE ROTINA:

```
source("I:\\MATERIAS_PROJETO E TESE DOUT_BIOMETIRA\\ANALISES_DOUT_2016
\\ANALISE_COX\\vis.gam.txt")
## OBTENDO O MAPA DE RISCO COM AS CURVAS DE NIVEL ##
obj1<-vis.gam(gamtp1,view=c("POINT_X","POINT_Y"),
plot.type="contour",color="cm",asp=1,xlab="latitude",
ylab="longitude",latlim=c(7448157,7485015),
longlim=c(269637,314110),bty="n",
main="Coorte1: 1980-1990",n.grid=200)
```

PARA RETIRAR PONTOS FORA DO PERÍMETRO

```
perim = expand.grid(x=obj1[[1]],y=obj1[[2]])
inside = inout(perim, coord_camp2)
obj1[[3]][!inside]<-NA
```

GERANDO O GRÁFICO

```
image(obj1[[1]],obj1[[2]],obj1[[3]],xlim=c(269637,314110),
ylim=c(7448157,7485015),zlim=range(obj1[[3]],na.rm=T),
col=gray(sort(100:300/300,decreasing =T)),xlab="latitude",
ylab="longitude",main="Coorte1: 1980-1990",bty="n")
contour(obj1[[1]],obj1[[2]],obj1[[3]],add=TRUE,zlim=
range(obj1[[3]],na.rm=T),nlevels=30,labcex=0.9)
#points(dados1$POINT_X, dados1$POINT_Y,pch=20,cex=0.2)
obj<-polygon(coord_camp2, lwd = 3)
## COORTE 2 - 1996-2000 ##
## CURVAS DE SOBREVIVENCIA PARA COORTE 2 ##
par(mfrow=c(2,4))
kmsexo<-survfit(Surv(tempo,censura)~sexo,data=dados2)
```

```

plot(kmsexo,col=1,lty=1:2,mark.time=F,bty="n",xlab="tempo (em dias)",
ylab="S(t)",main="Sexo")
legend(locator(1),legend=c("Feminino","Masculino"),col=1,
lty=c(1,2),bty="n",cex=1)
legend(locator(1),legend=c("valor p<0,01"),col=1,lty=0,cex=1,bty="n")
kmidade<-survfit(Surv(tempo,censura)~idade1,data=dados2)
plot(kmidade,lty=1:3,col=1,bty="n",xlab="tempo (em dias)",
ylab="S(t)",main="Idade")
legend(locator(1),col=1,lty=1:3,legend=c("<=30 anos",
"30-49 anos",">= 50 anos"),bty="n",cex=1)
legend(locator(1),legend=c("valor p<0,01"),col=1,lty=0,cex=1,bty="n")
kmesc<-survfit(Surv(tempo,censura)~escol,data=dados2)
plot(kmesc,lty=1:2,col=1,bty="n",xlab="tempo (em dias)",
ylab="S(t)",main="Escolaridade")
legend(locator(1),,lty=1:2,col=1,legend=c(">=8 anos","<8 anos"),
bty="n",cex=1)
legend(locator(1),legend=c("valor p<0,01"),col=1,lty=0,cex=1,bty="n")
kmhom<-survfit(Surv(tempo,censura)~hom_het,data=dados2)
plot(kmhom,lty=1:2,col=1,bty="n",xlab="tempo (em dias)",
ylab="S(t)",main="HSH/bissexual")
legend(locator(1),lty=1:2,col=1,legend=c("Nao","Sim"),bty="n",cex=1)
legend(locator(1),legend=c("valor p=0,47"),col=1,lty=0,cex=1,bty="n")
kmdrog<-survfit(Surv(tempo,censura)~droga_inj,data=dados2)
plot(kmdrog,lty=1:2,col=1,bty="n",xlab="tempo (em dias)",
ylab="S(t)",main="Droga injetavel")
legend(locator(1),lty=1:2,col=1,legend=c("Nao","Sim"),bty="n",cex=1)
legend(locator(1),legend=c("valor p<0,01"),col=1,lty=0,cex=1,bty="n")
kmcancer<-survfit(Surv(tempo,censura)~cancer,data=dados2)
plot(kmcancer,lty=1:2,col=1,bty="n",xlab="tempo (em dias)",

```

```

ylab="S(t)",main="C\\^{a}ncer")
legend(locator(1),lty=1:2,col=1,legend=c("Nao","Sim"),bty="n",cex=1)
legend(locator(1),legend=c("valor p=0,18"),col=1,lty=0,cex=1,bty="n")
kmtuber<-survfit(Surv(tempo,censura)~ANT_TUBERC,data=dados2)
plot(kmtuber,lty=1:2,col=1,bty="n",xlab="tempo (em dias)",
ylab="S(t)",main="Tuberculose")
legend(locator(1),lty=1:2,col=1,legend=c("Nao","Sim"),bty="n",cex=1)
legend(locator(1),legend=c("valor p<0,01"),col=1,lty=0,cex=1,bty="n")
kmimun<-survfit(Surv(tempo,censura)~imupress,data=dados2)
plot(kmimun,lty=1:2,col=1,bty="n",xlab="tempo (em dias)",ylab="S(t)",
main="Imunossupresao")
legend(locator(1),lty=1:2,col=1,legend=c("> 200","=< 200"),bty="n",cex=1)
legend(locator(1),legend=c("valor p<0,01"),col=1,lty=0,cex=1,bty="n")
## TESTE DE LOGRANK PARA COORTE 2 ##
survdifff(Surv(tempo,censura)~sexo,data=dados2)
survdifff(Surv(tempo,censura)~idade1,data=dados2)
survdifff(Surv(tempo,censura)~escol,data=dados2)
survdifff(Surv(tempo,censura)~droga_inj,data=dados2)
survdifff(Surv(tempo,censura)~hom_het,data=dados2)
survdifff(Surv(tempo,censura)~ANT_TUBERC,data=dados2)
survdifff(Surv(tempo,censura)~cancer,data=dados2)
survdifff(Surv(tempo,censura)~imupress,data=dados2)
### AJUSTE DO MODELO DE COX PARA COORTE 2 ##
mod2<-coxph(Surv(tempo,censura)~sexo+escol+hom_het+idade1+droga_inj+
cancer+ANT_TUBERC+imupress,data=dados2)
summary(mod2)
## AVALIANDO A SUPOSICAO DE RISCOS PROPORCIONAIS ##
cox.zph(mod2)
## AJUSTE MODELO GAM ##

```

```

gamtp2<-gam(tempo~sexo+escol+hom_het+idade1+droga_inj+cancer+ANT_TUBERC+
imupress+s(POINT_X,POINT_Y),family=cox.ph(),weights=censura,data=dados2)
### MODELO COM AS COVARIAVEIS SELECIONADAS ##
gamtp2<-gam(tempo~idade1+droga_inj+ANT_TUBERC+imupress+
s(POINT_X,POINT_Y),family=cox.ph(),weights=censura,data=dados2)
summary(gamtp2)
## AVALIANDO OS RESIDUOS MARTINGALE E DEVIANCE ####
resdev<-resid(gamtp2,type="deviance")
plot(resdev,xlab="Indice",ylab="Residuo de Deviance",
main="Coorte 2: 1996-2000")
abline(h=0,lty=3)
## MAPA DE RISCO ##
source("I:\\MATERIAS_PROJETO E TESE DOUT_BIOMETIRA\\
ANALISES_DOUT_2016\\ANALISE_COX\\vis.gam.txt")
obj1<-vis.gam(gamtp2,view=c("POINT_X","POINT_Y"),
              plot.type="contour",color="cm",asp=1,xlab="latitude",
              ylab="longitude",latlim=c(7448157,7485015),
              longlim=c(269637,314110),bty="n",
main="Coorte1",n.grid=200)
perim = expand.grid(x=obj1[[1]],y=obj1[[2]])
inside = inout(perim, coord_camp2)
obj1[[3]][!inside]<-NA
image(obj1[[1]],obj1[[2]],obj1[[3]],xlim=c(269637,314110),
ylim=c(7448157,7485015),
zlim=range(obj1[[3]],na.rm=T), col=gray(sort(100:300/300,
decreasing =T)),xlab="latitude",
        ylab="longitude",main="Coorte 2: 1996-2000",bty="n")
contour(obj1[[1]],obj1[[2]],obj1[[3]],add=TRUE,
zlim=range(obj1[[3]],na.rm=T),

```

```

nlevels=30,labcex=0.9)
#points(dados2$POINT_X, dados2$POINT_Y,pch=1,cex=0.0001,col=1)
obj<-polygon(coord_camp2, lwd = 3)
## COORTE 3 - 2001 - 2005 ##
## CURVAS DE SOBREVIVENCIA PARA COORTE 3 ##
par(mfrow=c(2,4))
kmsexo<-survfit(Surv(tempo,censura)~sexo,data=dados3)
plot(kmsexo,col=1,lty=1:2,mark.time=F,bty="n",xlab="tempo (em dias)",
ylab="S(t)",main="Sexo")
legend(locator(1),legend=c("Feminino","Masculino"),
col=1,lty=c(1,2),bty="n",cex=1)
legend(locator(1),legend=c("valor p<0,01"),col=1,lty=0,cex=1,bty="n")
kmidade<-survfit(Surv(tempo,censura)~idade1,data=dados3)
plot(kmidade,lty=1:3,col=1,bty="n",xlab="tempo (em dias)",
ylab="S(t)",main="Idade")
legend(locator(1),col=1,lty=1:3,
legend=c("<=30 anos","30-49 anos",">= 50 anos"),bty="n",cex=1)
legend(locator(1),legend=c("valor p<0,01"),col=1,lty=0,cex=1,bty="n")
kmesc<-survfit(Surv(tempo,censura)~escol,data=dados3)
plot(kmesc,lty=1:2,col=1,bty="n",xlab="tempo (em dias)",
ylab="S(t)",main="Escolaridade")
legend(locator(1),lty=1:2,col=1,legend=c(">=8 anos","<8 anos"),
bty="n",cex=1)
legend(locator(1),legend=c("valor p<0,01"),col=1,lty=0,cex=1,bty="n")
kmhom<-survfit(Surv(tempo,censura)~hom_het,data=dados3)
plot(kmhom,lty=1:2,col=1,bty="n",xlab="tempo (em dias)",
ylab="S(t)",main="HSH/bissexual")
legend(locator(1),lty=1:2,col=1,legend=c("Nao","Sim"),bty="n",cex=1)
legend(locator(1),legend=c("valor p=0,03"),col=1,lty=0,cex=1,bty="n")

```

```

kmdrog<-survfit(Surv(tempo,censura)~droga_inj,data=dados3)
plot(kmdrog,lty=1:2,col=1,bty="n",xlab="tempo (em dias)",
ylab="S(t)",main="Droga injetavel")
legend(locator(1),lty=1:2,col=1,legend=c("Nao","Sim"),bty="n",cex=1)
legend(locator(1),legend=c("valor p=0,16"),col=1,lty=0,cex=1,bty="n")
kmcancer<-survfit(Surv(tempo,censura)~cancer,data=dados3)
plot(kmcancer,lty=1:2,col=1,bty="n",xlab="tempo (em dias)",
ylab="S(t)",main="Cancer")
legend(locator(1),lty=1:2,col=1,legend=c("Nao","Sim"),bty="n",cex=1)
legend(locator(1),legend=c("valor p<0,01"),col=1,lty=0,cex=1,bty="n")
kmtuber<-survfit(Surv(tempo,censura)~ANT_TUBERC,data=dados3)
plot(kmtuber,lty=1:2,col=1,bty="n",xlab="tempo (em dias)",
ylab="S(t)",main="Tuberculose")
legend(locator(1),lty=1:2,col=1,legend=c("Nao","Sim"),bty="n",cex=1)
legend(locator(1),legend=c("valor p<0,01"),col=1,lty=0,cex=1,bty="n")
kmimun<-survfit(Surv(tempo,censura)~imupress,data=dados3)
plot(kmimun,lty=1:2,col=1,bty="n",xlab="tempo (em dias)",
ylab="S(t)",main="Imunossupresao")
legend(locator(1),lty=1:2,col=1,legend=c("> 200","=< 200"),bty="n",cex=1)
legend(locator(1),legend=c("valor p<0,01"),col=1,lty=0,cex=1,bty="n")
## TESTE DE LOGRANK PARA COORTE 3 ##
survdiff(Surv(tempo,censura)~sexo,data=dados3)
survdiff(Surv(tempo,censura)~idade1,data=dados3)
survdiff(Surv(tempo,censura)~escol,data=dados3)
survdiff(Surv(tempo,censura)~droga_inj,data=dados3)
survdiff(Surv(tempo,censura)~hom_het,data=dados3)
survdiff(Surv(tempo,censura)~ANT_TUBERC,data=dados3)
survdiff(Surv(tempo,censura)~cancer,data=dados3)
survdiff(Surv(tempo,censura)~imupress,data=dados3)

```

```

## AJUSTE DO MODELO DE COX ##
mod3<-coxph(Surv(tempo,censura)~sexo+escol+hom_het+idade1+droga_inj+
cancer+ANT_TUBERC+imupress,data=dados3)
summary(mod3)
## AVALIANDO A SUPOSICAO DE RISCOS PROPORCIONAIS ##
cox.zph(mod3)
VARIAVEL IMUNUSSUPRESSAO GRAVE FERE A SUPOSICAO DE RISCOS PROPORCIONAIS
### PROPOSTA ESTRATIFICAR O MODELO PELA VARIAVEL IMUNUSSUPRESSAO ####
### APOS A SELECAO DAS VARIAVEIS PELO METODO STEPWISE ##
mod3<-coxph(Surv(tempo,censura)~sexo+escol+strata(imupress),data=dados3)
summary(mod3)
(ressch3<-cox.zph(mod3,transform="identity"))
## AJUSTE DO MODELO GAM ESTRATIFICADO ##
gamtp3<-gam(cbind(tempo,imupress)~sexo+escol+hom_het+idade1+droga_inj+
cancer+ANT_TUBERC+s(POINT_X,POINT_Y),family=cox.ph(),
weights=censura,data=dados3)
## AJUSTE DO MODELO GAM ESTRATIFICADO SELECIONADO ##
gamtp3<-gam(cbind(tempo,imupress)~sexo+escol+s(POINT_X,POINT_Y),
family=cox.ph(),weights=censura,data=dados3)
## AVALIANDO OS RESIDUOS DE DEVIANCE ####
resdev<-resid(gamtp3,type="deviance")
plot(resdev,xlab="Indice",ylab="Residuo de Deviance",
main="Coorte 3: 2001-2005")
abline(h=0,lty=3)
## MAPA DE RISCO ##
source("I:\\MATERIAS_PROJETO E TESE DOUT_BIOMETIRA\\
ANALISES_DOUT_2016\\ANALISE_COX\\vis.gam.txt")
obj1<-vis.gam(gamtp3,view=c("POINT_X","POINT_Y"),
plot.type="contour",color="cm",asp=1,xlab="latitude",

```

```
        ylab="longitude",latlim=c(7448157,7485015),
        longlim=c(269637,314110),bty="n",
main="Coorte 3",n.grid=200)
perim = expand.grid(x=obj1[[1]],y=obj1[[2]])
inside = inout(perim, coord_camp2)
obj1[[3]][!inside]<-NA
image(obj1[[1]],obj1[[2]],obj1[[3]],xlim=c(269637,314110),
ylim=c(7448157,7485015),
zlim=range(obj1[[3]],na.rm=T),
col=gray(sort(100:300/300,decreasing =T)),xlab="latitude",
        ylab="longitude",main="Coorte 3: 2001-2005",bty="n")
contour(obj1[[1]],obj1[[2]],obj1[[3]],
add=TRUE,zlim=range(obj1[[3]],na.rm=T),
nlevels=30,labcex=0.9)
#points(dados3$POINT_X, dados3$POINT_Y,cex=0.0001,col=1)
obj<-polygon(coord_camp2, lwd = 3)
```