



APLICAÇÃO DE MODELOS NÃO PARAMÉTRICOS E DE RISCOS PROPORCIONAIS DE COX EM DADOS DE HCV/HIV

Beatriz Regina Brum¹, Omar Neves Pereira², Claudinei Mesquita da Silva³ e Isolde Previdelli⁴

¹Programa de Pós-graduação em Bioestatística - UEM

²Programa de Pós-graduação em Bioestatística - UEM

³Programa de Pós-graduação em Ciências da Saúde - UEM

⁴Departamento de Estatística - UEM

RESUMO

A coinfeção por vírus da Imunodeficiência Humana Adquirida e Hepatite C vem se agravando nas últimas décadas. Na corrida pela produção de medicamentos antirretrovirais potentes a ponto de inibir os prejuízos causados pelo HIV, negligenciou-se infecções com pelo vírus da hepatite C (HCV). Com o objetivo de verificar as chances dos indivíduos atingiram um sistema imunológico saudável, caracterizado pela contagem de CD4 acima de 600 células por mm^3 de sangue e a existência de diferença entre o grupo controle (indivíduos que possuem apenas HIV) e coinfectados (indivíduos com HCV/HIV), foi utilizado modelos não paramétricos como: Kaplan-Meier, Teste Log-rank e o modelo semiparamétrico de Riscos Proporcionais de Cox com intuito de estimar possível influência do grupo no decorrer do tempo até atingir a contagem de CD4 capaz de garantir um sistema imunológico saudável. Ao final da análise não foi verificada diferença significativa entre os grupos, acarretando em um efeito não significativo dos mesmos. Por fim concluiu-se que as chances de indivíduos do grupo controle e coinfectados atingirem um sistema imunológico considerado saudável são as mesmas.

Palavras chave: coinfeção pelo HCV/HCV, modelos não paramétricos e Modelo de Riscos Proporcionais de Cox.

1 INTRODUÇÃO

A preocupação da comunidade acadêmica em obter estimativas de taxa de cura ou de ocorrência do evento de interesse em determinado tempo é ponto crucial desse estudo. A disseminação de diversos tipos de enfermidade que acometem os seres vivos serve como motivação para questionamentos do tipo: “Qual a chance de atingir um evento em determinado tempo sabendo que até então ele não ocorreu?” ou “Qual seria a chance do evento ocorrer até um tempo determinado?”.

Esse trabalho buscou atender as necessidades do pesquisador em verificar possível diferença de chances entre um grupo de indivíduos infectados pelo vírus da Imunodeficiência Humana Adquirida (HIV) e outro coinfectados por HCV/HIV. Para isso foi utilizado modelos de análise de sobrevivência não paramétricos na análise descritiva dos dados e o modelo de semiparamétrico de Riscos Proporcionais de Cox para se possível, determinar a existência de efeito e chances dos indivíduos de diferentes grupos em atingir um sistema imunológico considerado saudável em determinado tempo.

A análise foi feita em um conjunto de dados referentes a pacientes do sudoeste do Paraná. Os dados são constituídos por uma amostra de 93 pacientes com HIV dos quais 32 coinfectados pelo vírus da hepatite C (HCV/HIV), ao final do estudo 58 indivíduos haviam sofrido o evento de interesse.

O trabalho está organizado da seguinte forma: a apresentação dos conceitos básicos de análise de sobrevivência são descritos na Seção 2, a aplicação dos dados são apresentados na Seção 3 e na quarta Seção estão as conclusões a cerca do estudo realizado.

2 METODOLOGIA

Entre os procedimentos utilizados para a análise descritiva dos dados estão técnicas não paramétricas, como o estimador Produto-Limite proposto por Kaplan-Meier e o teste Log-rank. Para finalizar a análise será utilizado o modelo semiparamétrico de Riscos Proporcionais de Cox.

2.1 Estimador de Kaplan-Meier

De acordo [3], entre os diversos métodos utilizados para estimar a função de sobrevivência o estimador Produto-Limite de Kaplan-Meier sem duvida nenhuma é o mais utilizado. O mesmo oferecendo uma primeira ideia do comportamento dos dados e permite observar o pressuposto de proporcionalidade exigidos pelo modelo de Cox. Para obtenção de estimativa da curva de sobrevivência, ordenam-se os tempos de evento de cada indivíduo da seguinte forma, $t_1 < t_2 < \dots < t_k$, sendo a sobrevivência estimada por:

$$S(t) = \prod_{i=1}^k \left(\frac{n_i - d_i}{n_i} \right), \quad (1)$$

em que n_i é o número de observações sobre risco no tempo t_i e d_i o número de falhas ou eventos no tempo [2].

2.2 Teste de Log-rank

O teste de Mantel-Haenszel (1959) é particularmente útil na comparação da sobrevivência [4]. A estatística de teste para dois estratos compara valores observados e esperados de ambos, sendo expressa como segue,

$$\text{Log-rank} = \frac{(O_1 - E_1)^2}{\text{Var}(O_1 - E_1)}, \quad (2)$$

A estatística segue uma distribuição qui-quadrado com um grau de liberdade, a hipótese testada é de igualdade entre curvas, dessa forma se o valor p computado for inferior ao nível de significância estipulado o teste será significativo.

2.3 Modelo de Riscos Proporcionais de Cox

Entre os modelos de regressão utilizados na análise de sobrevivência estão os modelos semiparamétricos, sua utilização abrange diversas áreas de conhecimento, principalmente no que diz respeito à estudos clínicos [3]. Um modelo muito conhecido é o proposto por Cox, o mesmo assume proporcionalidade de riscos e tem forma expressa por

$$\lambda(t|\mathbf{x}) = \lambda_0(t) \exp(\mathbf{x}\beta) \quad (3)$$

observe que se $\mathbf{x}\beta = 0$, então $\lambda(t|\mathbf{x}) = \lambda_0(t)$, onde $\mathbf{x} = (x_1, x_2, \dots, x_p)$, representa um vetor de covariáveis, $\beta = (\beta_1, \beta_2, \dots, \beta_p)$ é o vetor de parâmetros da regressão desconhecido que indica a magnitude do efeito das covariáveis no risco da ocorrência do evento. Para a estimação dos parâmetros o método de máxima verossimilhança usual não é adequado para o modelo semiparamétrico. A probabilidade da amostra para modelos paramétricos é escrito por

$$L(\beta) = \prod_{i=1}^n [\lambda_0(t_i) \exp(\mathbf{x}_i\beta)]^{\delta_i} (S_0(t_i))^{\exp(\mathbf{x}_i\beta)}, \quad (4)$$

o primeiro fator do produto corresponde à contribuição dos indivíduos que sofreram o evento de interesse e o segundo aos que foram censurados. Ao observar a expressão descrita em (4), verifica-se a presença da função de perturbação, $\lambda_0(t_i)$, de modo que o processo anterior de estimação não serve. O método de estimação utilizado para o modelo consiste em determinar a probabilidade do evento ocorrer no tempo t_i , dado o espaço amostral das indivíduos sob risco, denotado por $R(t_i)$.

A estimação é realizada a partir de verossimilhanças individuais, L_i , que representam a razão entre o risco do i -ésimo indivíduo sofrer o evento exatamente no instante anterior a t_i , pela soma dos demais indivíduos sob risco, a mesma é escrita como

$$L_i = \frac{\exp(\mathbf{x}_i\beta)}{\sum_{j \in R(t_i)} \exp(\mathbf{x}_j\beta)}. \quad (5)$$

Mais detalhes podem ser vistos em [3], [4] e [2].

3 RESULTADOS E DISCUSSÕES

O conjunto de dados é composto por uma amostra de 93 indivíduos portadores da Imunodeficiência Humana Adquirida (HIV), dos quais 61 pertencem ao grupo controle (HIV) e 32 ao grupo coinfestado (HCV/HIV). O acompanhamento foi realizado por 3600 dias, o evento de interesse foi o tempo até que a contagem de CD4 atingisse 600 células/mm³. O número de indivíduos que sofreram o evento foi de 58 e 35 foram censurados.

Por meio de Kaplan-Meier, é possível verificar que as curvas de sobrevivência não diferem bruscamente entre si e parecem ser proporcionais.

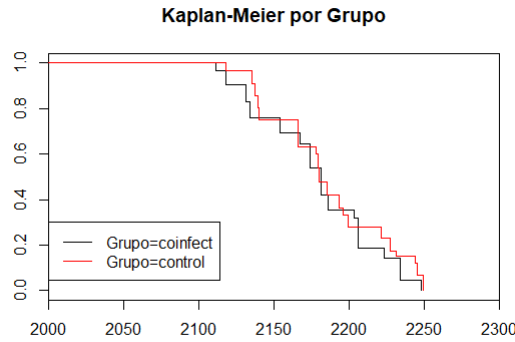


Figura 1: Estimação por Kaplan-Meier.

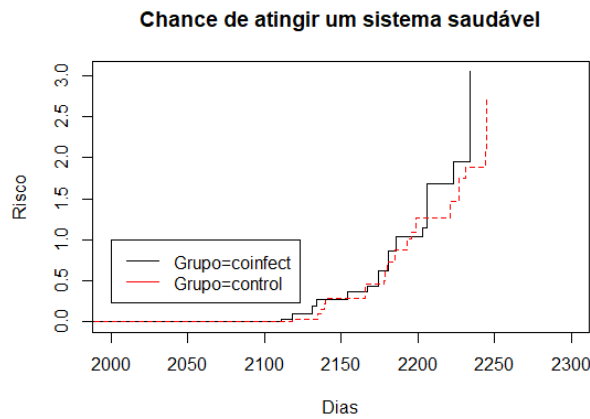


Figura 2: Curva de Risco.

O teste Log-rank computou um p valor de 0.137, que ao nível de significância de 10%, indica marginalmente não haver diferença entre o grupo controle e coinfestados.

Para uma ultima análise e abordando todas as metodologias apresentadas no trabalho, o modelo de Cox foi então utilizado para verificar se de fato a existência de influência do grupo na contagem do CD4 não é significativa, o modelo é expresso por

$$\lambda_i(t|K(t)) = \lambda_0(t) \exp \{ \gamma grupo_i \}, \quad (6)$$

em que γ é o efeito do grupo no tempo até o evento.

O resultado obtido verifica-se que o efeito associado ao grupo foi marginalmente não significativo a 10% de significância.

Tabela 1: Estimativa da influência do grupo no tempo até o evento

	coef	exp(coef)	se(coef)	z	p
Grupocontrol	-0,5165	0,5966	0,3357	-1,539	0,124

4 CONCLUSÃO

A estimativa atribuída ao grupo não foi significativa, corroborando com a análise descritiva realizada. Portanto os pacientes do grupo controle, como de coinfeção, apresentam praticamente as mesmas chances de atingir um sistema imunológico considerado saudável.

5 AGRADECIMENTOS

PBE, DCS, Fundação Araucária e Caixa Econômica Federal.

Referências

- [1] CARVALHO, M. S.; ANDREOZZI, V. L.; CODEÇO, C. T.; CAMPOS, D. P.; BARBOSA, M. T. S.; SHIMAKURA, S. E. Análise de Sobrevivência: teoria e aplicações em saúde. [S.l.]: SciELO-Editora FIOCRUZ, 2011.
- [2] COLLETT, D. Modelling survival data in medical research. [S.l.]: CRC press, 2015.
- [3] COLOSIMO, E. A.; GIOLO, S. R. Análise de sobrevivência aplicada. [S.l.]: Edgard Blücher, 2006. (ABE - Projeto Fisher).
- [4] LEE, E. T.; GO, O. T. Survival analysis in public health research. Annual review of public health, Annual Reviews 4139 El Camino Way, PO Box 10139, Palo Alto, CA 94303-0139, USA, v. 18, n. 1, p. 105–134, 1997.
- [5] KAPLAN, EDWARD L., AND PAUL MEIER. Nonparametric estimation from incomplete observations. Journal of the American statistical association 53.282 (1958): 457-481.
- [6] ZHAO, Q. and SUM, J. (2004), Generalized Log-Rank Test for Mixed Interval-Censored Failure Time Data, Statistics in Medicine, 23, 1621–1629.