

# Modelagem de doença pulmonar obstrutiva crônica através do modelo de regressão de Cox com interações

Cecilia Candolo<sup>1 3</sup>

Diego Marmorato Marino<sup>2</sup>

Kamilla Tays Marrara<sup>2</sup>

## 1 Introdução

Exacerbação é definida como aumentos agudos nos sintomas respiratórios que excedam a variação diária normal e que levem à necessidade de mudanças nas medicações regularmente em uso pelo paciente, caracterizando a exacerbação da doença pulmonar obstrutiva crônica (DPOC) (MARINO *et al.*, 2013). Pode ser caracterizada principalmente pela piora da dispnéia, implicando em prejuízos na capacidade de exercício e reduzindo não apenas as atividades físicas, mas também a qualidade e a expectativa de vida destes pacientes. Uma maneira de estudar a exacerbação é feita a partir da medição do tempo até a exacerbação do paciente, contado a partir do início do estudo, e tentar relacionar este tempo com medidas corporais do paciente além de medidas de dispneia, de força de prensão palmar e de tolerância ao esforço físico. Uma ocorrência comum em medidas de tempos até ocorrência de um evento é a não obtenção da medida, ou seja, o evento não ocorrer até o término do estudo. A medida para o paciente em questão é chamada de censura.

A modelagem de tempos até ocorrência de eventos, com presença de censura, em associação com covariáveis pode ser feita a partir de várias abordagens, sendo que o objetivo é modelar o efeito dessas covariáveis e para isso usa-se um modelo de regressão apropriado para dados censurados. Em análise de sobrevivência, existem duas classes de modelos propostas na literatura: os modelos paramétricos e os semiparamétricos. Os modelos paramétricos são mais eficientes, porém menos flexíveis do que os modelos semiparamétricos. A segunda classe de modelos, também denominada simplesmente de modelo de regressão de Cox, tem sido bastante utilizada em estudos clínicos. Além da flexibilidade, este modelo permite incorporar facilmente covariáveis dependentes do tempo, que ocorrem com frequência em várias áreas de aplicação (COLOSIMO e GIOLO, 2006).

Este trabalho aborda a análise dos dados de um estudo que teve como objetivo determinar os preditores de exacerbação, via o modelo de regressão de Cox, isoladamente ou em interação, em pacientes com doença pulmonar obstrutiva crônica (DPOC) do município de São Carlos- SP

---

<sup>1</sup>DEs - UFSCar. e-mail: [cecilia@ufscar.br](mailto:cecilia@ufscar.br)

<sup>3</sup>Agradecimento à ProEx/UFSCar pelo apoio.

<sup>2</sup>Dfisio - UFSCar.

em tratamento fisioterapêutico ao longo de seis meses no serviço de Fisioterapia Respiratória da UFSCar. A variável resposta foi o tempo até a ocorrência de exacerbação e as covariáveis foram índice de massa corpórea, massa magra, índice de massa magra, volume expiratório forçado no primeiro segundo, dispnéia, distância percorrida no teste de caminhada de seis minutos e força de preensão palmar. A partir dessa modelagem, procurou-se explorar a análise dos resultados na presença de interações significativas, através da construção de gráficos de funções de sobrevivência e de cálculos de riscos relativos, sempre explorando uma variável e mantendo as outras fixas, para explorar natureza interativa das covariáveis em estudo.

## 2 Material e métodos

Os dados utilizados neste trabalho são oriundos de um estudo longitudinal e prospectivo envolvendo pacientes com diagnóstico clínico de DPOC acompanhados por 06 meses. Este estudo foi realizado na Unidade Especial de Fisioterapia Respiratória e na Unidade Saúde Escola (USE) da Instituição, e foram avaliados 50 pacientes com diagnóstico clínico de DPOC, de ambos os gêneros, em tratamento nos referidos ambulatórios. O estudo foi aprovado pelo Comitê de Ética e Pesquisa da Instituição e os pacientes assinaram o Termo de Consentimento Livre e Esclarecido, atendendo à resolução 196/96 do Conselho Nacional de Saúde. Todos os pacientes realizaram as seguintes avaliações: coleta de dados antropométricos e demográficos, composição corporal por meio da massa magra (MM) e índice de massa magra (IMM); obstrução da via aérea, avaliada pelo VEF1, da dispneia pela escala modified Medical Research Council (mMRC), avaliação da força de preensão palmar e da tolerância aos esforços físicos através da distância percorrida (DP) no teste de caminhada de seis minutos (TC6). A exacerbação da DPOC foi caracterizada pela piora sustentada da condição de saúde dos pacientes que exigissem uma visita não agendada a um serviço de saúde como pronto socorro, unidade básica de saúde ou pneumologista responsável. As variáveis consideradas para a presente análise foram codificadas como: tempo até a ocorrência de exacerbação (tempo, com cens=1 para observados e 0 para censurados) em semanas, índice de massa corpórea (IMC), massa magra (MassaMagra), índice de massa magra (IMM); volume expiratório forçado no primeiro segundo (VEF1), dispnéia (MRC), distância percorrida no teste de caminhada de seis minutos (TC6) e força de preensão palmar (Palmar).

Para a apresentação do modelo de regressão de Cox, considere  $p$  covariáveis e seja  $\mathbf{x}$  o vetor de componentes  $\mathbf{x} = (x_1, \dots, x_p)$  medidas para cada indivíduo. O modelo de regressão de Cox é da seguinte forma

$$\lambda(t) = \lambda_0(t) \exp(\mathbf{x}'\boldsymbol{\beta}) = \lambda_0(t) \exp(x_1\beta_1 + \dots + x_p\beta_p),$$

em que  $\lambda(t)$  é a taxa de falha no tempo  $t$ ,  $\lambda_0(t)$  é a taxa de falha base no tempo  $t$  e  $\boldsymbol{\beta}$  é o vetor dos parâmetros associados às covariáveis. Esse modelo é composto pelo produto de dois

componentes, um não paramétrico e outro paramétrico, em que o componente não paramétrico  $\lambda_0(t)$  não é especificado. Já o componente paramétrico é representado por  $\exp(\mathbf{x}'\beta)$ . Muitas vezes este modelo também é chamado de modelo de riscos proporcionais, pois a suposição básica para o uso do modelo de regressão de Cox é que as taxas de falha sejam proporcionais, ou de forma equivalente, as taxas de falha acumuladas sejam também proporcionais, isto é, a razão das funções de taxas de falhas para os indivíduos  $i$  e  $j$  é representada por

$$\frac{\lambda_i(t)}{\lambda_j(t)} = \frac{\lambda_0(t) \exp(\mathbf{x}'_i\beta)}{\lambda_0(t) \exp(\mathbf{x}'_j\beta)} = \exp(\mathbf{x}'_i\beta - \mathbf{x}'_j\beta).$$

A estimação dos parâmetros do modelo é feita pelo método da verossimilhança parcial. Para os estimadores podem ser calculados seus desvios padrões, construídos testes e intervalos de confiança para os parâmetros, para a qualidade do ajuste, para a verificação da suposição de riscos proporcionais e análise de diagnóstico, ver COLOSIMO e GIOLO (2006). O maior interesse na modelagem geralmente é a estimação dos coeficientes de regressão, entretanto, as funções relacionadas a  $\lambda_0(t)$  são também importantes no modelo de Cox. Estas funções referem-se basicamente à função de taxa de falha acumulada  $\Lambda_0(t) = \int_0^t \lambda_0(u) du$  e à correspondente função de sobrevivência de base  $S_0(t) = \exp(-\Lambda_0(t))$ . Vale lembrar também que, a função de sobrevivência  $[S_0(t)]^{\exp(\mathbf{x}'\beta)}$  é útil quando se deseja concluir a análise em termos de percentis associados a grupos de indivíduos. Com relação à interpretação dos parâmetros estimados, pode-se observar que o efeito das covariáveis no modelo é o de acelerar ou desacelerar a função de risco. A função de risco instantânea no tempo  $t$ , ou risco relativo, é obtida pela razão das funções de taxas de falhas para os indivíduos  $i$  e  $j$ , que têm os mesmos valores para as covariáveis com exceção da  $l$ -ésima. Temos  $\lambda_i(t)/\lambda_j(t) = \exp[\beta_l(x_{il} - x_{jl})]$ .

### 3 Resultados e discussões

A análise dos dados foi realizada utilizando os pacotes R (R, 2013) e SAS (SAS, 2011). O R foi utilizado para a análise descritiva, ajustes, verificação de suposições, análise de resíduos e seleção de variáveis. Foi também utilizado para a construção de gráficos das funções de sobrevivência estimadas e gráficos de interações. O modelo final foi ajustado no SAS, versão 9.3, que proporciona ferramentas facilitadoras para o cálculos de razões de risco e seus intervalos de 95% de confiança, que ajudam a explorar a compreensão dos resultados na presença de interações significativas. Depois de feitos os ajustes iniciais e verificadas suposições do modelo, a seleção de variáveis selecionou como melhor modelo aquele com as variáveis IMC, MassaMagra e TC6 com as interações TC6 com IMC, TC6 com MassaMagra e IMC com MassaMagra. O modelo foi significativo ( $p = 0,0011$ ) e as interações foram também significativas, sendo a interação entre TC6 e IMC ( $p = 0,0101$ ) e TC6 e Massa Magra ( $p = 0,0133$ ) significativas a 5% e a interação entre IMC e Massa Magra foi significativa a 10% ( $p = 0,0602$ ). A presença dessas interações significativas dificultam a interpretação do resultados, pois o efeito de cada

covariável sobre o tempo até a exacerbação não poderá ser considerado sozinho, é necessário considerar os níveis das demais. Para explorar o entendimento das interações foram calculadas as funções de sobrevivência e riscos relativos estimados para alguns valores de uma covariável, mantendo as outras fixas. Os valores utilizados para covariáveis, com base em informações da literatura e nos valores das 1os. e 3os. quartis da amostra, foram, para TC6: 340,357,500; para IMC: 20,23,25 e para Massa Magra (MM): 43,46,50. As estimativas resultantes foram em grande quantidade. Aqui serão apresentadas apenas parte dos resultados. Por exemplo, a Figura 1 apresenta alguns gráficos das funções de sobrevivência estimadas variando os valores de Massa Magra. Neste gráfico observamos que para algumas combinações de valores, as funções são muito parecidas (no centro, TC6=357 e IMC=23) e para outras apresentam bastante disparidade (embaixo, à esquerda, TC6=500 e IMC=20).

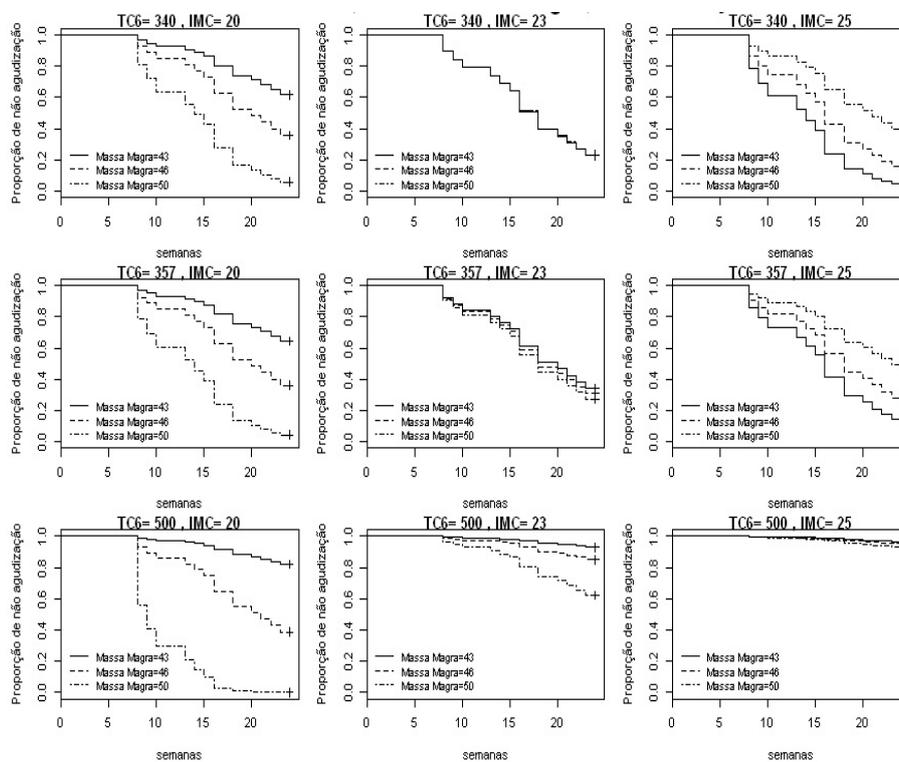


Figura 1: Funções de sobrevivência estimadas, ou proporção de não agudização (exacerbação), variando os valores de Massa Magra.

O comando HAZARDRATIO do procedimento PHREG do SAS (versão 9.2 em diante), permite o cálculo de riscos relativos e seus intervalos de confiança, com variações de diferença de unidades (UNITS) para as comparações. Por exemplo, o comando HAZARDRATIO MassaMagra / at (TC6=340 357 500 IMC=20 23 25) UNITS=3; produz como um dos resultados (ao todo são 9 resultados, para esse caso):

MassaMagra Unit=3 At TC6=340 IMC=20 2.180 1.093 4.346,

para o qual a interpretação é que o risco de agudização (exacerbação) em pacientes com massa magra 46 é, em média, aproximadamente 2 vezes (2.180) maior que o risco de agudização para

pacientes com massa magra 43, mantendo fixas TC6 em 340 e IMC em 20. Foram também construídos gráficos de riscos relativos para diferenças de uma unidade, como, por exemplo, na Figura 2, que apresenta estes riscos calculados para MassaMagra, ao longo de TC6, para diferentes valores de IMC (à esquerda) e ao longo de IMC, para diferentes valores de TC6 (à direita).

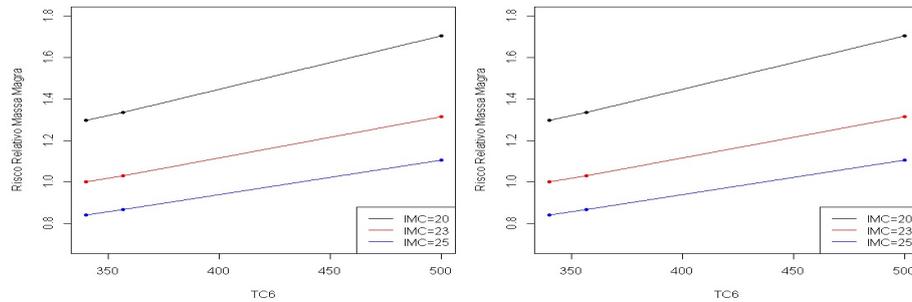


Figura 2: Riscos relativos para diferenças de uma unidade calculados para MassaMagra, ao longo de TC6, com diferentes valores de IMC (à esquerda) e ao longo de IMC, com diferentes valores de TC6 (à direita).

## 4 Conclusões

A partir dos cálculos feitos para várias combinações diferentes de valores das covariáveis envolvidas nas interações significativas, foi possível uma compreensão mais aprofundada, para os profissionais da área da saúde, da natureza interativa das covariáveis em estudo e sua relação com o risco de exacerbação para os pacientes com doença pulmonar obstrutiva crônica.

## Referências

- [1] COLOSIMO, E. A.; GIOLO, S. R. **Análise de Sobrevida Aplicada**. São Paulo: Editora Edgard Blücher Ltda, ABE-Projeto Fisher, 2006, 367p.
- [2] MARINO, D. M.; MARRARA, K. T.; ARCURI, J. F.; CANDOLO, C.; JAMAMI, M.; PIRES DI LORENZO, V. A. Determinação dos preditores de exacerbação nos pacientes com DPOC em tratamento fisioterapêutico estudo longitudinal. Submetido em **Revista Brasileira de Fisioterapia**. 2013.
- [3] R Development Core Team **R: A language and environment for statistical computing**. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0, URL <http://www.R-project.org/>, 2013.
- [4] SAS Institute Inc. Cary, NC, USA. 2011.