



ASPECTOS ASSOCIADOS AO ÓBITO POR CÂNCER DE MAMA VIA MODELOS DE REGRESSÃO BINÁRIA

Ricardo Rasmussen Petterle¹, Roberta Gurnacki de Wallau Petterle², Gabriel Pizzatto Rudey Crovador³, Celso Schmitt Amorim⁴ e Eduardo Schunemann Júnior⁵

¹⁻⁵ Universidade Federal do Paraná-UFPR.

RESUMO

O principal objetivo deste estudo foi analisar os aspectos associados ao óbito por câncer de mama das pacientes atendidas no Hospital de Clínicas da Universidade Federal do Paraná no período de janeiro de 2003 a dezembro de 2012. Foram comparados sete modelos de regressão binária, sendo quatro que não dependem de parâmetros extras (*logit*, *probit*, *complemento log-log* e *cauchit*) e três que dependem de parâmetros extras (*Aranda-Ordaz*, *Weibull* e *Prentice*). Pelos critérios de informação de *Akaike* e *Bayesiano* o modelo de regressão logística mostrou-se mais adequado para análise dos dados. Os resultados indicaram que as pacientes que fizeram quimioterapia, não realizaram cirurgia e estão gestantes foram as mais propensas ao óbito.

Palavras chave: Câncer de mama; regressão binária.

1 INTRODUÇÃO

O câncer é um grupo de doenças caracterizado pelo crescimento descontrolado de células anormais. Essas células multiplicam-se rápido e de forma agressiva, ocasionando tumores malignos que podem se espalhar para outros órgãos e regiões do corpo (American Cancer Society, 2015).

O câncer, de uma forma geral, é a segunda causa de mortalidade em mulheres em idade reprodutiva, sendo que o câncer de mama especificamente é uma das neoplasias malignas mais diagnosticadas durante a gestação (Amant *et al.*, 2012).

O câncer de mama é um problema de saúde pública, a nível mundial. No Brasil, para o ano de 2016, estima-se mais de 57 mil novos casos de câncer de mama (Inca, 2016).

Seu diagnóstico pode ser feito tanto pelo auto-exame das mamas, como pela mamografia ou ultra sonografia, sendo confirmado pela biópsia (Trufelli *et al.*, 2008). A confirmação da doença traz sofrimento psicológico às pacientes, devido às perdas sociais, emocionais e materiais, causadas pelo tratamento (Silva, 2008).

O principal objetivo deste estudo foi analisar os aspectos associados ao óbito por câncer de mama das pacientes atendidas no Hospital de Clínicas da Universidade Federal do Paraná no período de janeiro de 2003 a dezembro de 2012.

2 MATERIAL e MÉTODOS

O conjunto de dados provém do Registro Hospitalar de Câncer do Hospital de Clínicas da Universidade Federal do Paraná (HC-UFPR). Foram selecionados todos os casos de câncer de mama diagnosticados e tratados no hospital no período de janeiro de 2003 a dezembro de 2012. O número total de casos no período foi de 667. Deste total, foram selecionadas as pacientes em idade fértil, até 49 anos de idade, totalizando uma amostra de 248 prontuários para análise. A variável resposta deste estudo é binária (1-óbito; 0-não óbito). As covariáveis são: idade, gestação, tipo histológico, estadiamento clínico, cirurgia, radioterapia, quimioterapia e hormonioterapia.

Os resultados da pesquisa foram analisados através do *software* R (R Core Team, 2016), versão 3.3.1, e só foram considerados resultados relevantes aqueles com $p < 0,05$. As covariáveis foram selecionadas através do método *forward*.

Para os ajustes dos modelos de regressão binária foram consideradas as seguintes funções de ligação:

- *logit*: $g(\mu_i) = \ln \left(\frac{\mu_i}{1 - \mu_i} \right)$;
- *probit*: $g(\mu_i) = \Phi(\mu_i)$, em que $\Phi(\cdot)$ denota a função de distribuição da Normal padrão;
- *complemento log-log*: $g(\mu_i) = \ln(-\ln(1 - \mu_i))$;
- *cauchit*: $g(\mu_i) = \tan \left(\pi \left(\mu_i - \frac{1}{2} \right) \right)$;
- *Aranda-Ordaz*: $g(\mu_i) = \ln \left\{ \frac{(1 - \mu_i)^{-\delta} - 1}{\delta} \right\}$, com $\delta > 0$;
- *Weibull*: $g(\mu_i) = [-\ln(1 - \mu_i)]^{1/\delta}$, com $\delta > 0$;
- *Prentice*: $g(\mu_i)$ é obtida a partir do logaritmo da função de distribuição acumulada da densidade \mathcal{F} com dois parâmetros, δ_1 e δ_2 .

3 RESULTADOS

Na Tabela 1 são apresentados quatro modelos de regressão binária que não dependem de parâmetros extras (*logit*, *probit*, *complemento log-log* e *cauchit*) e três modelos que dependem de parâmetros extras (*Aranda-Ordaz*, *Weibull* e *Prentice*). Pelos critérios de informação de *Akaike* (AIC) e *Bayesiano* (BIC), no qual o melhor modelo é aquele que apresenta o menor valor, foi selecionado o modelo de regressão binária com ligação *logit*, popularmente conhecido como modelo de regressão logística.

Tabela 1: Comparação entre sete modelos de regressão binária.

Modelos	NP	Log-verossimilhança	AIC	BIC
logit	4	-124,44	256,89	270,94
probit	4	-124,45	256,91	270,96
complemento log-log	4	-125,43	258,86	272,91
cauchit	4	-125,02	258,05	272,11
Aranda-Ordaz	5	-123,82	257,65	275,21
Weibull	5	-124,42	258,85	276,42
Prentice	6	-123,39	258,79	279,87

NP = número de parâmetros

A partir do modelo selecionado, estimou-se os coeficientes de regressão. A Tabela 2 mostra as covariáveis do modelo final, estimativas dos parâmetros, erros-padrão e *p*-valores. As categorias de referência são para o grupo de pacientes que não fez quimioterapia, não fez cirurgia e não estão gestantes. Pode-se observar que todas as covariáveis apresentam significância, pois *p*-valor < 0,05.

Tabela 2: Estimativas dos parâmetros, erros-padrão e *p*-valores do modelo de regressão logística.

Covariável	Coeficiente	Erro-Padrão	<i>p</i> -valor
β_0 : constante	-1,571	0,442	0,00038
β_1 : quimioterapia	1,781	0,471	0,00016
β_2 : cirurgia	-1,371	0,362	0,00015
β_3 : gestante	1,352	0,667	0,04255

Por exemplo, a chance de uma paciente que fez quimioterapia vir a óbito é aproximadamente 6 vezes maior em relação a daquela que não fez quimioterapia, pois o cálculo da razão de chances se dá pela seguinte expressão $\exp\{1,781\} \approx 6$.

CONCLUSÕES

Este estudo teve por objetivo analisar os aspectos associados ao óbito por câncer de mama das pacientes atendidas no Hospital de Clínicas da Universidade Federal do Paraná no período de janeiro de 2003 a dezembro de 2012. Foram comparados sete modelos de regressão para dados binários. Utilizando os critérios de informação de

Akaike e Bayesiano selecionou-se o modelo de regressão logística para análise dos dados.

Os resultados indicaram que as pacientes que fizeram quimioterapia, não realizaram cirurgia e estão gestantes foram as mais propensas ao óbito.

Referências

- [1] AMERICAN CANCER SOCIETY. Cancer facts & figures 2015. Atlanta: American Cancer Society; 2015. Disponível em: <http://www.cancer.org/acs/groups/content/@editorial/documents/document/acspc-044552.pdf>. Acesso em: 30 out. 2016.
- [2] AMANT, F; LOIBL, S.; NEVEN, P.; CALSTEREN, K.V. Breast Cancer in Pregnancy. *The Lancet, Bélgica*. v.379, n.9815, p.570-579, 2012.
- [3] INCA. Instituto Nacional de Câncer. Tipos de cancer - Mama. Disponível em: <http://www2.inca.gov.br/wps/wcm/connect/tiposdecancer/site/home/mama>. Acesso em: 30 out. 2016.
- [4] R DEVELOPMENT CORE TEAM. 2016. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria.
- [5] TRUFELLI, D. C; et al. Análise do atraso no diagnóstico e tratamento do câncer de mama em um hospital público. *Revista da Associação Médica Brasileira* (1992. Impresso), v.54, n.1, p.72-76, 2008.
- [6] SILVA, L. C. Câncer de mama e sofrimento psicológico: aspectos relacionados ao feminino. *Psicologia em Estudo* (Impresso), v.13, n.2, p.231-237, 2008.