

INTRODUÇÃO AOS MODELOS DE FRAGILIDADES APLICADOS A DADOS DE LEUCEMIA LINFOBLÁSTICA

Roseane de Alcântara Costa (1); Manoel Joaquim Isidro (1); Tiago Almeida de Oliveira (4)

Universidade Estadual da Paraíba, rose_ane_costa@hotmail.com

Universidade Estadual da Paraíba, joakym_@hotmail.com

Universidade Estadual da Paraíba, tadolive@cct.uepb.edu.br

RESUMO: Dentre as análises estatísticas, a análise de sobrevivência é uma das mais utilizadas na área da medicina, dada sua utilidade em estimar o risco dos pacientes curar-se ou morrer ao longo do estudo. Neste trabalho, objetivou-se produzir um material prático e claro sobre a análise de dados de sobrevivência, que possa vir a auxiliar aos que desejarem utilizar os modelos de fragilidades. Uma característica marcante da análise de sobrevivência é a presença de censura nos dados, isto é, refere-se ao acompanhamento do paciente ser interrompido, por diversas razões tais como, desistência voluntária, mudança de cidade, morte do paciente entre outras. A análise de sobrevivência modela o tempo até a ocorrência do evento de interesse e incorpora a informação das censuras, ou seja, utiliza o tempo até a censura dos pacientes que participaram do estudo e não falharam. Os dados foram obtidos do Institute for Health & Society da Universidade de Wisconsin (Medical College of Wisconsin), o banco de dados é constituído de um total de 137 pacientes (38 LLA, 99 LMA) os quais apresentam uma distinção entre os tipos de câncer, leucemia mieloide aguda (LMA) e leucemia linfoblástica aguda (LLA), os pacientes foram tratados em quatro hospitais. Na realização deste trabalho será estimado a curva de sobrevivência não paramétrica de Kaplan-Meier para os 3 grupos de pacientes com diferentes tipos de leucemia, em seguida utilizaremos o teste log-rank para investigar se existe diferença significativa entre as curvas de sobrevivência, após ajustarmos as distribuições paramétricas e, por fim, aplicaremos o modelo de Cox e os modelos de fragilidades Gama e Log-Normal por pacientes. A análise será realizada utilizando o software R.

Palavras-chave: Leucemia Linfoblástica, Modelos de Fragilidades, Análise de Sobrevivência.

INTRODUÇÃO

A análise de sobrevivência é formada por um conjunto de técnicas estatísticas para analisar dados, que consistem nos tempos até a ocorrência de um evento de interesse, comumente conhecido como tempo de sobrevivência. São exemplos de casos de análise de sobrevivência o tempo até que um aluno possa concluir sua graduação assim como a morte do paciente após o diagnóstico de uma certa doença. Um fato que caracteriza dados de sobrevivência é a possibilidade da presença de censura, que é a ocorrência da observação parcial da resposta de interesse. A variável resposta foi o tempo em dias até a morte do paciente ou até o término do acompanhamento.

Leucemia Linfoblástica Aguda (LLA) é um câncer que se origina de um grupo de células precursoras dos linfócitos. Os linfócitos são glóbulos brancos que defendem o corpo contra infecções. A medula óssea cria inúmeras células subdesenvolvidas conhecidas como blastos. A LLA é resultado de um dano genético adquirido (não herdado) no DNA de um grupo de células

(glóbulos brancos) na medula óssea. As células doentes substituem a medula óssea normal e os efeitos são o crescimento incontrolável e o acúmulo de linfoblasto (linfócitos imaturos) que perdem a capacidade de funcionar como células sanguíneas normais, levando a um bloqueio ou diminuição na produção de glóbulos vermelhos, plaquetas e glóbulos brancos na medula óssea. Após o diagnóstico, o tratamento da LLA é dividido em duas partes: terapia de indução e terapia pós-indução ou transplante de medula óssea. A quimioterapia utiliza medicamentos anticancerígenos para destruir as células tumorais, dependendo do tipo e do estado da leucemia, e pode ser utilizada sozinha ou combinada com a radioterapia (PEDROSA; LINS, 2002).

Depois de um transplante de medula óssea a recuperação é um processo complexo e lento. O prognóstico para recuperação pode depender de fatores de risco conhecidos, no momento do transplante, como a idade do paciente e/ou idade do doador e o sexo, a fase inicial da doença, o tempo a partir do diagnóstico para o transplante, etc. O prognóstico final pode mudar após o transplante do paciente, pois a ocorrência de eventos em momentos aleatórios durante a recuperação, como o desenvolvimento de doenças aguda ou crônica do enxerto-versus-hospedeiro (GVHD). O transplante pode ser considerado uma falha quando houver a recaída da doença (recaída de leucemia no paciente), ou quando o paciente morre enquanto está em remissão da doença (morte relacionada com o tratamento). O Transplante de Medula Óssea configura-se como um grande problema de saúde pública tanto nos países desenvolvidos como nos países em desenvolvimento. Há muitas causas diferentes para a doença, que variam de radiação de produtos químicos a causas genéticas, tem o objetivo de destruir as células leucêmicas, para que a medula óssea volte a produzir células normais.

Diante do exposto o uso de estatística mais especificamente da análise de sobrevivência pode contribuir para explicar os fatores de risco para o tempo de sobrevivência de pacientes transplantados e o uso das diversas técnicas da análise de sobrevivência permitirá a escolha do melhor método para o tipo de estudo em questão.

Objetivo geral deste trabalho é de utilizar o modelo de Fragilidade de Cox nos dados de leucemia linfoblástica. Dentro do qual ajustaremos os métodos clássicos e introdutórios da análise de sobrevivência e determinaremos a fragilidade de pacientes com leucemia linfoblástica.

METODOLOGIA

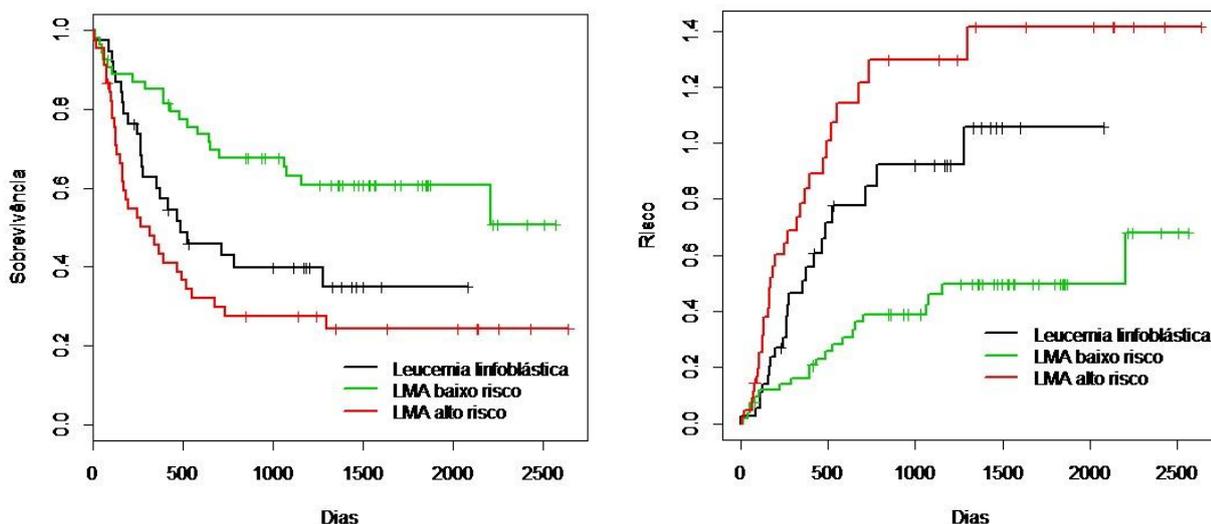
Nesse trabalho foi feita uma revisão teórica sobre a análise de sobrevivência e suas principais ferramentas estatísticas, onde foi estudado o tempo até a ocorrência da morte dos pacientes após o transplante de medula óssea. Também foram estudadas as características que constituem o banco de dados, realizando uma aplicação dessas ferramentas estatísticas em um banco de dados. Os resultados obtidos serão apresentados neste trabalho. Os dados foram obtidos do Institute for Health & Society da universidade de Wisconsin (Medical College of Wisconsin), estes dados estão descritos detalhadamente em Klein e Moeschberger (2003). O banco de dados é constituído de um total de 137 pacientes (38 LLA, 99 LMA) os quais possuem uma distinção entre os tipos de câncer, leucemia mieloide aguda (LMA) e leucemia linfoblástica aguda (LLA) foram tratados em um de quatro hospitais avaliados, as variáveis estudadas foram: Z_1 : Idade do paciente; Z_2 : Idade do doador; Z_3 : Sexo do paciente; Z_4 : Sexo do doador; Z_5 : Citomegalovírus do paciente; Z_6 : Citomegalovírus do doador; Z_7 : Tempo de espera em dias para o transplante de medula óssea; Z_8 : FAB; Z_9 : Hospital; Z_{10} : MTX medicamento administrados nos pacientes. gg_1 : Grupo de risco LLA;

gg₂: Grupo LMA baixo risco; gg₃: Grupo de risco LMA alto risco; *a* = GVHD agudo; *c* =GVHD crônico; *p* = Plaquetas em níveis normais(recuperação das plaquetas). Aplicou-se o método não paramétrico de Kaplan-Meier e o teste log-rank por grupos às distribuições paramétricas de Weibull, Exponencial, Log-Normal e Gama, para estimar a função de sobrevivência e por fim aplicou-se o modelo tradicional de Cox e os modelos de Fragilidades Gama e Log-normal por pacientes. A aplicação das técnicas e métodos de análise de sobrevivência foi possível com auxílio do ambiente R (R CORE TEAM, 2014) e do pacote survival (THERNEAU, 2014).

RESULTADOS E DISCUSSÕES

A seguir demonstram-se os principais resultados obtidos a partir de uma análise realizada com o auxílio do software R.

A curva de sobrevivência de Kaplan-Meier foi aplicada ao banco de dados, considerando os três grupos constituídos por (grupo 1 - LLA; grupo 2 - LMA baixo risco e o grupo 3 - LMA alto risco).

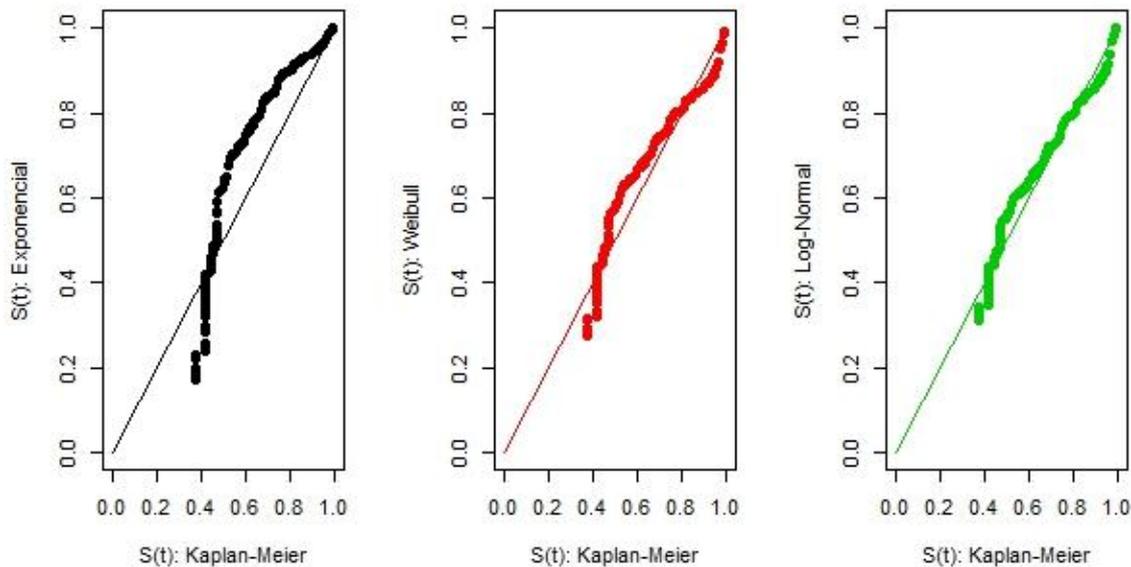


De acordo com o gráfico de sobrevivência, observa-se que os pacientes dos três grupos sobreviveram mais que 1500 dias após transplante de medula óssea, percebe-se também que aproximadamente 60% dos pacientes do grupo (LMA baixo risco) sobreviveram mais de 1500 dias, enquanto os pacientes do grupo (LMA alto risco) aproximadamente 30% sobreviveram 1500 dias. Os pacientes do grupo (LMA alto risco) possuem a menor sobrevivência entre os grupos. Já gráfico do risco, mostra que o risco de óbito dos pacientes aumenta com o transcorrer do tempo. Este comportamento revela um efeito gradual da gravidade da doença em cada grupo, percebe-se também que os pacientes do grupo 3 (LMA alto risco) apresenta um risco maior que os demais grupos de vir á óbito, o que era esperado.

Após a aplicação do método de Kaplan-Meier, efetuou-se a comparação entre as curvas por meio do teste log-rank, para investigar se existe diferença entre as curvas, resultando numa estatística $\chi^2 = 16,3$ para 2 graus de liberdade e um p-valor igual a 0,000294, ou

seja, indicando que de fato existe diferença significativa entre as três curvas, ou seja, os pacientes possuem tempos de vida diferentes entre os grupos.

Ajustando as distribuições paramétricas tem-se:



De acordo com os gráficos, é possível ver que a distribuição Exponencial parece não ser adequado para explicar esses dados. Por outro lado, as distribuições de Weibull e Log-Normal acompanham mais de perto a reta $y = x$, indicando ser uma dessas distribuições, possivelmente, adequada para explicar os dados sobre estudo. Para selecionarmos a distribuição, fez-se necessário aplicar o teste da razão de verossimilhança e o critério de Akaike.

Modelo	$\log(L(\theta))$	TRV	Valor p	AIC (menor melhor)
Gama generalizado	-632.2	-	-	-
Weibull	-627.6	$2(627.6-632.2)=-9.2$	1	1275.34
Log-normal	-623.7	$2(623.7-632.2)=-17.0$	1	1268.42

De acordo com os resultados obtidos na tabela acima, tendo como base o valor do AIC, tem-se que a distribuição Log-Normal e a que melhor se ajustou aos dados em estudo.

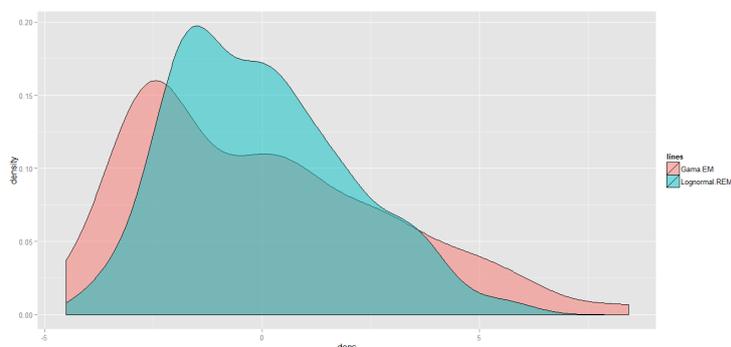
Modelo de Riscos Proporcionalis de Cox

variáveis	exp(coef)	exp(-coef)	Limite inferior	Limite superior
p	0.3786	2.6413	0.2028	0.7067
c	0.4254	2.3506	0.2614	0.6923
z8	1.5954	0.6268	0.8930	2.8502
z9	0.7215	1.3861	0.5300	0.9821
z10	2.1616	0.4626	1.1157	4.1878
gg1	0.9318	1.0732	0.4718	1.8404
gg2	0.4358	2.2945	0.2398	0.7920

De acordo com o modelo ajustado de Riscos proporcionais de Cox, observamos que variável doença do hospedeiro (c) se mostrou como fator protetor: pacientes sem a doença do hospedeiro têm risco maior de ir a óbito do que os pacientes que tem a doença do hospedeiro. Este efeito protetor pode ser interpretado como um efeito indireto. A variável FAB (Z_8) apresentou-se como um fator de risco, enquanto que a variável hospital (Z_9) atuou como fator protetor, assim como a variável MTX (Z_{10}) se mostrou como um fator de risco. O modelo apresentou um poder explicativo absoluto de 32,2%. A probabilidade de concordância estimada pelo modelo teve alto valor discriminatório ou preditivo (76,8%).

Fragilidade-LogNormal	exp(coef)	exp(-coef)	Limite inferior	Limite superior
p	0.08425	11.8693	0.026780	0.2651
c	0.23011	4.3457	0.104061	0.5089
z8	2.40000	0.4167	0.967981	5.9505
z9	0.63312	1.5795	0.416213	0.9631
z10	3.36885	0.2968	1.242339	9.1353
gg1	0.75457	1.3253	0.255363	2.2297
gg2	0.23945	4.1763	0.094842	0.6045

Na tabela acima têm se o modelo de Fragilidade Log-Normal, que utilizou 21 interações do algoritmo de Newton-Raphson para estimar a variância dos efeitos aleatórios, o modelo apresentou uma concordância estimada de alto valor discriminatório ou preditivo (95,9%), ou seja, o modelo foi bem estimado. A variância estimada dos efeitos aleatórios do modelo foi 2,5249. A fragilidade foi significativa. Os pacientes do grupo gg_2 apresentou uma fragilidade que os pacientes do gg_1 .



De acordo com gráfico das fragilidades dos pacientes, percebe-se que em ambos os modelos é possível destacar três picos, justamente os três grupos considerados no estudo (LLA, LMA baixo risco e alto risco).

CONCLUSÕES

O objetivo deste estudo foi utilizar o modelo tradicional de Cox e os modelos de Fragilidades, aplicado em dados de pacientes transplantados de leucemia linfoblástica, ajustou-se os métodos clássicos e introdutórios da análise de sobrevivência.

Inicialmente, foram ajustadas as curvas de Kaplan-Meier aos dados de pacientes transplantados de medula óssea, as curvas indicaram existir diferença entre os grupos, em seguida aplicou-se o teste log-rank e confirmou-se que as curvas apresentaram diferença entre si. Evidenciou-se que os pacientes do grupo LMA alto risco, foram os que apresentaram a maior probabilidade de óbito em relação aos demais grupos.

A distribuição de probabilidade que foi usada para explicar os dados da amostra foi a Log-Normal selecionada pelas técnicas gráficas e pelos testes da razão de verossimilhança e o critério de Akaike. O modelo de Riscos Proporcionais de Cox e o modelo de fragilidades Log-Normal foram adequados para estimar o risco dos pacientes com Leucemia. O modelo de fragilidade Log-Normal revelou que a fragilidade dos pacientes do grupo LMA de baixo risco é menor que os pacientes do grupo LLA, indicando que os pacientes do grupo LMA tem uma resistência melhor à doença.

A partir das técnicas estatísticas usadas nesse trabalho, evidenciou-se que análise de sobrevivência é uma importante ferramenta na área da saúde desde que todos os critérios que cada técnica estatística possui sejam seguidos de forma correta, ajudando a melhor entender o comportamento e quais características afetam os pacientes sobre risco.

REFERÊNCIAS

- COLOSIMO, E.A.; GIOLO, S.R. **Análise de Sobrevivência Aplicada**. 1ª edição. São Paulo: Editora Edgard Blucher, 2006.
- COLOSIMO, E.A. **A Generalized Log-Normal Model for Grouped Survival Data**. Communications in Statistics. Theory and Methods, v. 39, p. 2659-2666, 2010.
- COX, D. R. Regression models and life-tables. **Journal of the Royal Statistical Society. Series B (Methodological)**, p. 187-220, 1972.
- COX, D. R.; BARNDORFF-NIELSEN, O. E. **Inference and asymptotics**. CRC Press, 51 1994.
- KLEIN, John P.; MOESCHBERGER, Melvin L. **Survival analysis: techniques for censored and truncated data**. Springer Science & Business Media, 2003.
- PEDROSA, F.; LINS, M. **Leucemia linfóide aguda: uma doença curável**. Revista brasileira saúde materna infantil, v. 2, n. 1, p. 63-68, 2002.
- THERNEAU T (2014). Survival: A Package for Survival Analysis in S. R package version 2.37-7, URL <http://CRAN.R-project.org/package=survival>.