

Análise de sobrevivência através de um método não-paramétrico

Survival analysis using a nonparametric method

Edith Maria Formentini Chiesa^I, Anaelena Bragança Moraes^{II},
 Maria Helena Rigão^{III}

RESUMO

Partindo do interesse em estudar as técnicas de Análise de Sobrevivência, efetuou-se a revisão de literatura direcionada a estimação de curvas de sobrevivência através do estimador não-paramétrico de Kaplan-Meier e testes para comparação entre curvas: Log-rank e Qui-quadrado. Analisou-se os dados de pacientes com câncer no aparelho digestivo do Hospital Universitário de Santa Maria. Na comparação dos pacientes, quanto ao tipo de cirurgia, não houve diferença significativa entre as curvas de sobrevivência. Quanto aos estágios da doença (I, II, III e IV), houve diferença significativa entre as curvas de sobrevivência do estágio I em relação ao estágio III e IV e estágio II em relação ao IV. Quanto ao tipo de tratamento curativo e paliativo, houve diferença significativa entre as curvas. Conclui-se que esses resultados de probabilidade de sobrevivência podem auxiliar o profissional da área da saúde na decisão pelo melhor tratamento para os seus pacientes.

Palavras-chave: Análise de sobrevivência, Kaplan Meier

ABSTRACT

Based on the interest in studying Survival Analysis techniques, a literature review was conducted directed to the estimation of survival curves using the non-parametric Kaplan-Meier estimator and tests for comparison between log-rank and chi-square curves. . Data from cancer patients in the digestive tract of the University Hospital of Santa Maria were analyzed. In the comparison of patients, regarding the type of surgery, there was no significant difference between survival curves. As for the disease stages (I, II, III and IV), there was a significant difference between the survival curves of stage I in relation to stage III and IV and stage II in relation to IV. Regarding the type of curative and palliative treatment, there was a significant difference between the curves. It is concluded that these survival probability results may assist the healthcare professional in deciding the best treatment for their patients.

Keyword: Survival analysis, Kaplan Meier

^I Universidade Federal de Santa Maria, Brasil; e-mail: edithchiesa@gmail.com;

^{II} Universidade Federal de Santa Maria, Brasil. e-mail: anaelena@smail.ufsm.br;

^{III} Universidade Federal de Santa Maria, Brasil; e-mail: hrigao@yahoo.com.br



1 INTRODUÇÃO

Em alguns estudos de seguimento temporal, na área da saúde, o pesquisador tem como objetivo avaliar o tempo que decorre até a ocorrência de um evento de interesse (falha). Nesse tipo de estudo os sujeitos podem permanecer em observação, no estudo, por tempos distintos. Alguns sujeitos saem do estudo pela ocorrência do evento de interesse, porém outros perdem o seguimento por diferentes motivos, ou ainda, o estudo se encerra. Para contemplar essas situações (censuras), foram desenvolvidos métodos estatísticos denominados de Análise de Sobrevivência (COLOSIMO, 2001).

Se o seguimento dos sujeitos (tempo de sobrevivência) acontecesse até a ocorrência da falha de todos os sujeitos do estudo, ou seja, do evento de interesse, poderiam ser utilizadas as técnicas clássicas de análise estatística. Como nesse caso, o estudo termina sem que todos os sujeitos tenham falhado (sujeitos censurados) é necessário utilizar a Análise de Sobrevivência, que considera a ocorrência de dados censurados (COLOSIMO, 2001; COX, 1984). Com isso as técnicas chamadas de Análise de Sobrevivência são amplamente utilizadas na área da saúde, constando de técnicas não-paramétricas e técnicas semi-paramétricas.

Para estimar as probabilidades de sobrevivência pode ser utilizado o estimador de Kaplan-Meier, objetivo desse trabalho, e o estimador de Cox, sendo que este último faz parte das técnicas semi-paramétricas.

Na Análise de Sobrevivência, como em qualquer análise estatística, a etapa inicial recomendada consiste na descrição dos dados, sendo posteriormente utilizadas técnicas estatísticas complementares, entre elas, o estimador de Kaplan-Meier, para a função de sobrevivência, e os testes log-rank e Qui-quadrado para a comparação entre curvas.

Neste contexto, a variável de interesse é o tempo de sobrevivência, que consiste no tempo decorrido desde a entrada do indivíduo no estudo, ou tempo inicial até a ocorrência do evento de interesse (falha) ou censura. (ECHEVESTRE, 1997)

Portanto, esse trabalho teve como objetivo abordar as técnicas não-paramétricas de Análise de Sobrevivência numa abordagem teórica, bem como realizar

uma aplicação utilizando dados reais obtidos no Hospital Universitário de Santa Maria (HUSM). Os pacientes do estudo eram portadores de câncer no aparelho digestivo e foram submetidos a procedimento cirúrgico, conforme a avaliação do caso.

2 MATERIAL E MÉTODOS

Em Análise de Sobrevivência a variável de interesse é o tempo de sobrevivência, que consiste no tempo decorrido desde a entrada do indivíduo no estudo ou tempo inicial até a ocorrência do evento de interesse (falha) ou de censura.

É importante, em estudos de sobrevivência, definir de forma clara e precisa o que vem a ser a falha, tal como morte ou recidiva.

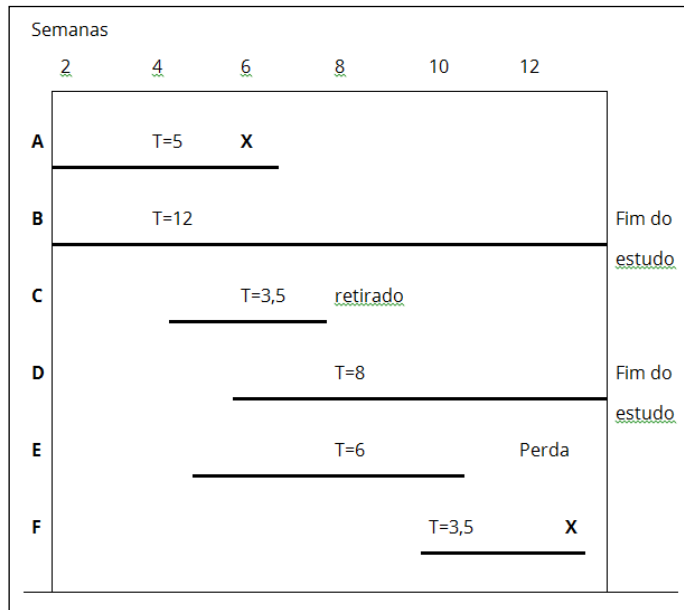
Uma característica decorrente de estudos clínicos de sobrevivência é a presença de observações incompletas ou censuradas. Podem ocorrer três tipos de censura:

- Censura do tipo I: um indivíduo não apresenta o evento de interesse até o fim do estudo, ou seja, o estudo termina e o evento não ocorre;
- Censura do tipo II: um indivíduo é perdido durante o período de estudo;
- Censura aleatória: um indivíduo é retirado do estudo, devido ao óbito por outra causa que não o evento de interesse, ou alguma outra razão. Estas situações estão ilustradas na Figura 1.

Na Figura 1 descreve-se a experiência de vários indivíduos acompanhados no tempo. Um X denota um indivíduo que apresentou o evento de interesse.

- O indivíduo A - apresentou o evento interesse;
- O indivíduo B - ocorreu a censura do tipo I;
- O indivíduo C - ocorreu a censura aleatória;
- O indivíduo D - ocorreu a censura tipo I;
- O indivíduo E - ocorreu a censura tipo II;
- O indivíduo F - apresentou o evento interesse.

Figura 1 - Descrição do acompanhamento de indivíduos no tempo



A variável aleatória contínua não-negativa t , que representa o tempo de sobrevivência, é usualmente especificada pela sua função de sobrevivência ou pela função risco (LEE, 1992; LATORRE, 1996).

a) Função densidade de probabilidade, $f(t)$:

A função densidade de probabilidade, também denominada taxa de falha não condicional, corresponde à probabilidade de que um indivíduo falhe num curto intervalo de tempo entre t e $t + \Delta t$, quando Δt tende a zero.

$$f(t) = \lim_{\Delta t \rightarrow 0} \frac{1}{\Delta t} P(t \leq T \leq t + \Delta t) \quad (1)$$

onde:

Δt = intervalo de tempo;

t = tempo de sobrevida específico;

T = variável aleatória contínua (tempo de sobrevida);

P = probabilidade.

cuja distribuição acumulada é a função $F(t)$, definida por:

$$F(t) = P(T \leq t) = \int_0^t f(s) ds \quad (2)$$

b) Função de sobrevivência (HOSMER; LEMESHOW, 1998)

A função de sobrevivência, é a proporção de indivíduos vivos a um tempo fixo t , ou além deste, definida como a probabilidade de uma observação não falhar até um certo tempo t , ou seja, a probabilidade de uma observação sobreviver ao tempo t . Em termos probabilísticos, isto é escrito como:

$$S(t) = P(T > t) \quad (3)$$

onde:

T = tempo de falha de uma determinada observação;

$S(t)$ = proporção de indivíduos sobreviventes no tempo t , ou além deste, se expresso como uma proporção (SOARES; COLOSIMO, 1995).

c) Função de taxa de falha ou de risco (KLEINBAUM, 1996)

Refere-se à probabilidade condicional de falhar durante um intervalo de tempo muito pequeno, dado que o indivíduo sobreviveu até o início do intervalo.

$$h(t) = \lim_{\Delta t \rightarrow 0} \frac{1}{\Delta t} P(t \leq T \leq t + \Delta t / T \geq t)$$

A função de sobrevivência $S(t)$ relaciona-se à função de densidade $f(t)$, por meio da distribuição $F(t)$, conforme a equação (3):

$$S(t) = P(T > t) = 1 - P(T \leq t) = 1 - F(t)$$

$$\int_0^t f(t) = 1 - S(t)$$

$$f(t) = 0 - \frac{ds(t)}{dt} = -S'(t)$$

A função de riscos $h(t)$ relaciona-se com as demais por meio de:

$$h(t) = \frac{f(t)}{S(t)} = \frac{-S'(t)}{S(t)} = \frac{-u'}{u}$$

$$h(t) = \frac{-S'(t)}{S(t)}$$

A função de riscos acumulada $H(t)$ é calculada por:

$$H(t) = \int_0^t \frac{-S'(t)}{S(t)}$$

$$H(t) = -\int_0^t \frac{u'}{u} = \frac{du}{u} = -\ln(u) = -\ln S(t)$$

$$S(t) = e^{-H(t)} = e^{-\int_0^t h(s)ds}$$

O estimador de Kaplan-Meier é um estimador não-paramétrico para a função de sobrevivência. Ele foi proposto por Kaplan e Meier (1958) e é também chamado de estimador limite-produto.

Supondo que n indivíduos sejam observados e, além disso, que o tempo de falha é observado em um dos indivíduos a k tempos distintos $t_1 < t_2 < \dots < t_k$. Sendo d_j o número de falhas no tempo t_j . Os outros w_j indivíduos são observações censuradas. Se uma censura e uma falha ocorrem ao mesmo tempo, considera-se que o tempo verdadeiro de falha para o indivíduo censurado é maior do que o tempo de censura observado (FISHER; BELLE, 1995). Sendo n_j o número de indivíduos sob risco de falha no tempo t_j , isto é,

$$n_j = n - d_{j-1} - w_{j-1}$$

A estimativa da curva de sobrevivência, por meio do estimador de Kaplan-Meier (HOSMER; LEMESHOW, 1998), é:

$$\hat{S}(t) = 1, \text{ para } t < t_1,$$

$$\hat{S}(t) = \prod_{j/t_j < t} \frac{n_j - d_j}{n_j}, t_j \leq t < t_{j+1} (j < k),$$

onde:

n_j = números de indivíduos sob risco;

d_j = número de falhas no tempo t_j ;

k = tempos distintos;

A fórmula para o erro padrão do estimador é conhecida como a fórmula de Greenwood:

$$SE(\hat{S}(t)) = \hat{S}(t) \sqrt{\sum_{j/t_j < t} \frac{d_j}{n_j(n_j - d_j)}}, \text{ para } t_j \leq t < t_{j+1}.$$

O tempo de sobrevivência mediano corresponde ao valor 0,5 da função de sobrevivência $\hat{S}(t)$ e representa o tempo no qual a metade dos pacientes sobreviveu, ou seja, não falhou.

Para a comparação entre duas curvas de sobrevivência utilizou-se o teste log-rank e o teste do Qui-quadrado para comparar três curvas ou mais.

O teste log-rank é apropriado quando a razão das funções de risco dos grupos a serem comparados é, aproximadamente, constante.

A estatística do teste log-rank é a diferença entre o número observado de falhas em cada grupo e o correspondente número esperado de falhas sob a hipótese nula.

Hipóteses:

H_0 = as curvas não apresentam diferença significativa;

H_1 = as curvas apresentam diferença significativa.

Suponha que no tempo t_j acontecem d_j falhas e n_j indivíduos estão sob risco em um tempo imediatamente inferior a t_j na amostra combinada e, respectivamente, d_{ij} e n_{ij} na amostra i ; $i = 1, 2$ e $j = 1, \dots, k$. Em cada tempo de falha t_j , os dados podem ser dispostos em forma de uma tabela de contingência 2x2 com d_{ij} falhas e $n_{ij} - d_{ij}$ censuras na coluna i , de acordo com a Tabela 1.

Tabela 1 – Tabela de contingência gerada no tempo t_j .

	Grupos		Total
	1	2	
Falha	d_{1j}	d_{2j}	d_j
Não Falha	$n_{1j} - d_{1j}$	$n_{2j} - d_{2j}$	$n_j - d_j$
Total	n_{1j}	n_{2j}	n_j

O valor esperado do número de falhas d_{1j} é $e_{1j} = n_{1j}d_j/n_j$ ou seja, se não houver diferença entre as duas populações no tempo t_j , o número total de falhas, d_j pode ser

dividido entre as duas amostras de acordo com a razão entre o número de indivíduos sob risco em cada amostra e o número total sob risco. A variância de d_{1j} é

$$V_j = \sum_j \frac{n_{1j}n_{2j}d_j(n_j - d_j)}{n_j^2(n_j - 1)} \quad (4)$$

Então, a estatística $(d_{1j} - e_{1j})$ tem média zero e variância V_j . Se as k tabelas de contingência forem independentes, um teste aproximado para a igualdade das duas funções de sobrevivência, pode ser baseado na estatística:

$$\chi^2 = \frac{(\sum d_{1j} - \sum e_{1j})^2}{\text{Var}(d_{1j} - e_{1j})} \quad (5)$$

que é a estatística do teste log-rank, tendo distribuição Qui-quadrado, com 1 grau de liberdade, para grandes amostras (ARMITAGE; BERRY, 1994).

Uma aproximação da estatística log-rank, pode ser calculada utilizando valores observados e esperados para cada grupo, sem o cálculo da variância da equação (4). A fórmula aproximada é a do Qui-quadrado (equação 5), com 1 grau de liberdade (HOSMER; LEMESHOW, 1998)

Para a comparação de mais de duas curvas de sobrevivência, pode-se utilizar o teste do Qui-quadrado, que considera somente os valores observados e esperados das falhas, sem necessitar do cálculo das variâncias e das covariâncias, necessários para o teste log-rank. A fórmula equivale a do teste clássico do Qui-quadrado, com $r-1$ graus de liberdade, sendo que a vantagem da utilização desta aproximação é a redução do esforço computacional (KLEINBAUM, 1996). O teste log-rank também pode ser utilizado para comparar três ou mais curvas de sobrevida (ARMITAGE; BERRY, 1994).

Os dados utilizados para a Análise de Sobrevivência são de pacientes portadores de câncer no aparelho digestivo e foram submetidos a procedimento cirúrgico, conforme a avaliação do caso, sendo atendidos no Hospital Universitário de Santa Maria (HUSM). O evento de interesse do estudo foi a morte dos pacientes por complicações relacionadas à doença. A morte de um paciente que não fosse causada pela doença foi considerada como sendo uma "perda", sendo este paciente censurado. As variáveis de interesse foram: idade, tipo de cirurgia, data da cirurgia, estadiamento da doença, tipo

de tratamento cirúrgico e data de óbito ou de censura. Os tipos de cirurgias do estudo foram: gastrectomia total, gastrectomia subtotal e irressecável. Quanto ao estadiamento, os pacientes foram classificados conforme os estágios I, II, III ou IV da doença. O tipo de tratamento cirúrgico foi classificado em: curativo ou paliativo. Uma informação relevante é que, quanto maior o estágio da doença maior a probabilidade do tratamento ser paliativo.

As probabilidades cumulativas de sobrevivência foram obtidas por meio do estimador não-paramétrico de Kaplan-Meier, para os pacientes do grupo geral, por tipo de cirurgia, por estadiamento e por tipo de tratamento cirúrgico, calculando-se também o erro padrão destas estimativas, sendo obtido o tempo de sobrevivência mediano, que é uma medida complementar da análise. As curvas de sobrevivência foram comparadas por meio do teste log-rank, no caso de duas curvas, ou do teste do Qui-quadrado, no caso de mais de duas curvas.

Para a análise dos dados foram utilizados uma planilha eletrônica do tipo Excel e o aplicativo computacional Statistica (Statsoft) (STATISTICA, 1998). Não foram considerados os aspectos éticos, pois na época (ano de 2002) esses aspectos não eram observados. Uma limitação deste trabalho foi a não discussão dos resultados de sobrevivência dos pacientes com artigos na área médica, pois não fazia parte dos objetivos do estudo.

3 RESULTADOS

A amostra estudada foi composta por 69 pacientes que apresentavam câncer no aparelho digestivo e foram tratados com procedimento cirúrgico. O período de admissão dos pacientes no grupo foi de outubro de 1991 a abril de 1999, e o seguimento desses pacientes encerrou-se em junho de 1999, sendo o tempo de seguimento máximo de, aproximadamente, 7,5 anos e o tempo mediano de sobrevivência de um ano.

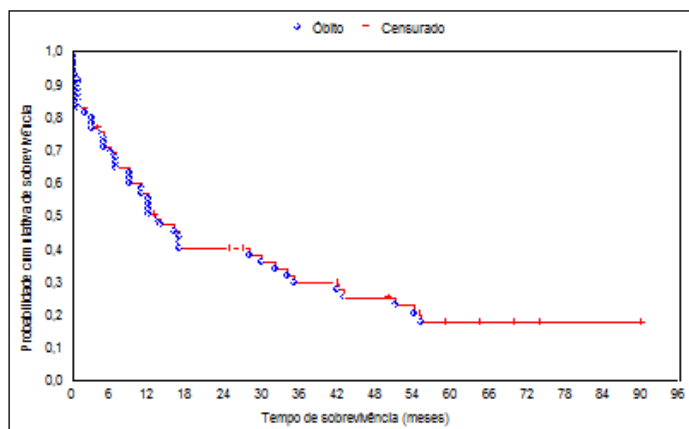
A data da cirurgia foi adotada como tempo inicial, ou seja, o instante da entrada dos pacientes no grupo de estudo. Como tempo final foi considerado a data da ocorrência de falha (óbito do paciente) ou a data do término do estudo (paciente censurado). O grupo de pacientes apresentou média de idade de 60 anos, mínima de

23 anos, máxima de 80 anos e desvio padrão de 12,3 anos.

Dos 69 pacientes estudados, 49 (71,0%) morreram devido a causas relacionadas com a doença em estudo, sendo que os 20 (29,0%) pacientes restantes foram censurados, ou seja, até o final do estudo estavam vivos ou morreram por outra causa. Considerando os 69 pacientes, a probabilidade cumulativa de sobrevivência estimada inicia em 1 (ou 100%) no tempo de sobrevivência zero e decai para 17,6% a partir do tempo de sobrevivência de 55 meses, permanecendo este mesmo valor até 90 meses (7,5 anos), pois neste período não ocorreu mais nenhum óbito. O tempo mediano de sobrevivência para o grupo geral foi de 1 ano, isto significa que 50% dos pacientes sobreviveram até este tempo, sendo que os outros 50%, sobreviveram mais de 1 ano. (Figura 1)

A função de sobrevida acumulada foi estimada pelo método de Kaplan-Meier e a curva de sobrevida dos 69 pacientes pode ser observada na Figura 1.

Figura 1 - Curva de sobrevida global



Na Tabela 2 são apresentados os números de óbitos, de sobreviventes e o total de pacientes de acordo com as categorias das variáveis consideradas. Quanto ao tipo de cirurgia, 20 pacientes submeteram-se a gastrectomia total, sendo que destes 80,0% morreram e 20,0% sobreviveram. Dos 36 pacientes submetidos à gastrectomia subtotal, 63,9% morreram e dos 13 pacientes que realizaram cirurgia irresssecável, 76,9% morreram. A probabilidade de um indivíduo do grupo que realizou gastrectomia

total, sobreviver até um ano foi de 58,2%, sendo o tempo mediano de sobrevivência, aproximadamente, 13 meses.

Os pacientes submetidos à gastrectomia subtotal apresentaram probabilidade de sobreviver mais do que um ano igual a 57,1%, sendo esta sobrevida semelhante ao grupo gastrectomia total, onde o tempo mediano de sobrevivência foi de 1,4 anos (17 meses). No grupo irressecável, a sobrevida em um ano foi de 17,6%, bem inferior aos outros tipos de cirurgia, como também o tempo mediano foi inferior. Isto significa que neste grupo, 50% dos indivíduos sobreviveram somente até 3 meses, enquanto que nos outros tipos de cirurgia, gastrectomia total e gastrectomia subtotal, os tempos medianos foram maiores, ou seja, 13,5 meses e 17,0 meses, respectivamente.

O resultado do teste para a comparação destas curvas não foi significativo, pois para 2 graus de liberdade e $p = 0,094$. Pode-se concluir que não existe diferença significativa entre as curvas de sobrevivência, por tipo de cirurgia.

Tabela 2 – Distribuição dos pacientes

	Número de óbitos (%)	Número de sobreviventes (%)	Número de pacientes (%)
Cirurgia			
Gastrectomia total	16 (80,0)	4 (20,0)	20 (29,0)
Gastrectomia subtotal	23 (63,9)	13 (36,1)	36 (52,2)
Irressecável	10 (76,9)	3 (23,1)	13 (18,8)
Total	49 (71,0)	20 (29,0)	69 (100,0)
Estágio			
I	2 (18,2)	9 (81,8)	11 (15,9)
II	4 (57,1)	3 (42,9)	7 (10,1)
III	20 (83,3)	4 (16,7)	24 (34,9)
IV	13 (100,0)	0 (0,0)	13 (18,8)
Sem classificação	10 (71,4)	4 (28,6)	14 (20,3)
Total	49 (71,0)	20 (29,0)	69 (100,0)
Tratamento			
Curativo	26 (61,9)	16 (38,1)	42 (60,9)
Paliativo	13 (100,0)	0 (0,0)	13 (18,8)
Sem classificação	10 (71,4)	4 (28,6)	14 (20,3)
Total	49 (71,0)	20 (29,0)	69 (100,0)

Estágio Curativo = I+ II+III; Estágio Paliativo = IV

No estágio I, estágio inicial da doença, foram classificados 11 pacientes, sendo que destes 18,2% morreram. Dos 7 pacientes no estágio II 57,1% morreram. No estágio

III, foram classificados 24 pacientes sendo que 83,3% morreram. No estágio IV, o estágio mais avançado da doença, dos 13 pacientes, nenhum sobreviveu. Dos pacientes, 14 não foram classificados segundo o estadiamento, sendo que 71,4% morreram.

Para o grupo de pacientes classificados no estágio I da doença, observou-se uma pequena probabilidade de óbito, sendo a probabilidade de sobrevivência em um ano de 77,8%.

Para os pacientes classificados no estágio II da doença, observou-se que a probabilidade de sobrevida em 16 meses foi de 71,4%, enquanto que os pacientes classificados no estágio III apresentaram uma probabilidade de sobrevida em 13 meses de 44,9%.

No estágio IV da doença, a probabilidade de sobrevida em 12 meses foi pequena, sendo de 7,7%, resultado esperado, em função de que este é o estágio mais avançado da doença.

Os tempos medianos de sobrevivência para os estágios da doença foram, aproximadamente, 30 meses para o estágio II, 10 meses para o estágio III, e 4 meses para o estágio IV. Os tempos medianos do estágio I e dos pacientes, sem classificação, não foram calculados, porque a probabilidade de sobrevivência não atingiu 50%.

Para a comparação das curvas de sobrevivência, por estadiamento, foi utilizado o teste do Qui-quadrado, obtendo-se o valor de $\chi_c^2 = 19,027$, com 3 graus de liberdade e $p = 0,0003$, constatando-se a existência de diferença significativa entre as curvas de sobrevivência.

Para identificar quais as curvas que diferiam significativamente, foi utilizado o teste log-rank, obtendo-se os seguintes resultados: a curva de sobrevivência do estágio I difere da curva do estágio III ($p = 0,003$) e da curva do estágio IV ($p = 0,001$) e, a curva do estágio II, difere da curva do estágio IV ($p = 0,004$), sendo as demais comparações, não-significativas, ao nível de 5%.

Quanto ao tipo de tratamento cirúrgico 42 pacientes realizaram cirurgia curativa, sendo que destes 61,9% morreram. Dos 13 pacientes que realizaram cirurgia paliativa, nenhum sobreviveu. Dos pacientes estudados 14 não tiveram classificação, sendo que 71,4% morreram.

Para o grupo de pacientes que realizou tratamento cirúrgico curativo, observou-se que a probabilidade de sobrevida em um ano foi de 62,6%, e para o grupo que realizou tratamento cirúrgico paliativo, esta probabilidade foi de 7,7%. Os tempos medianos de sobrevivência para estes dois grupos foram de aproximadamente: 17,0 meses e 4,0 meses, respectivamente. Pode-se observar que o tempo mediano do grupo, que foi submetido à cirurgia curativa foi bem maior do que o grupo submetido a cirurgia paliativa, dando uma indicação inicial de que, quanto maior o tempo mediano, maior será a probabilidade de sobrevivência do grupo, devendo ser constatado por meio da construção e comparação das curvas de sobrevivência.

Para a comparação das curvas foi utilizado o teste log-rank, sendo o resultado do teste significativo, $p = 0,003$. Concluiu-se que existe diferença significativa entre as curvas de sobrevivência por tipo de tratamento curativo e paliativo, sendo que os pacientes submetidos ao tratamento curativo apresentaram maior probabilidade de sobrevida. Isto pode ser observado no decaimento das curvas, pois a curva do grupo de pacientes submetidos ao tratamento cirúrgico paliativo apresenta decaimento mais intenso do que a do grupo submetido ao tratamento curativo.

À época, o objetivo deste trabalho, que resultou em uma monografia de especialização, era realizar um estudo sobre as técnicas de Análise de Sobrevivência utilizando os métodos não-paramétricos, com uma aplicação a dados da área médica. Com isso, não foi objetivo discutir os resultados com outras publicações da área médica sobre câncer do aparelho digestivo, e sim utilizar a técnica em dados reais utilizando um aplicativo computacional, sem foco na discussão.

4 CONCLUSÕES

Pode-se verificar que por meio do cálculo do tempo de sobrevivência mediano, bem como do comportamento da curva obtida pelo estimador não-paramétrico de Kaplan-Meier, em relação ao seu decaimento, pode-se ter uma boa idéia de qual grupo apresenta maior ou menor probabilidade de sobrevivência. Mas, para verificar a significância da diferença das curvas de sobrevivência uma das opções é utilizar o teste log-rank para duas amostras ou o teste do Qui-quadrado para mais de duas amostras.

Também, de acordo com as referências consultadas pode-se utilizar o teste log-rank para mais de duas amostras, sendo o teste do Qui-quadrado um teste aproximado para situações de duas ou mais amostras.

O tipo de cirurgia não apresentou diferença significativa ($p > 0,05$) na sobrevida dos pacientes, quanto aos diferentes procedimentos cirúrgicos a que estes foram submetidos. Entretanto, o estágio da doença apresentou diferença significativa ($p \leq 0,05$), quanto a sobrevida dos pacientes, sendo que, quanto mais avançado o estágio da doença menor a probabilidade de sobrevivência, ou seja, maior o risco de óbito do paciente.

Quanto ao tipo de tratamento cirúrgico curativo ou paliativo, sabe-se que há uma associação com o estágio da doença, pois os pacientes que estavam no estágio I, II e III receberam o tratamento curativo, já os pacientes que estavam no estágio IV receberam o tratamento paliativo por estarem em uma fase muito avançada da doença. Na comparação das curvas, como já era esperada, pela relação entre estas covariáveis, a sobrevivência dos pacientes foi, significativamente menor ($p \leq 0,05$) no grupo que recebeu o tratamento curativo.

Com este trabalho, pode-se avaliar a importância da Análise de Sobrevivência quando as observações são censuradas e sua aplicabilidade em pesquisas médicas, principalmente, no estudo de doenças com risco de óbito. Com estes resultados o profissional da área da saúde pode ter mais segurança na tomada de decisão sobre a escolha do tratamento, considerando aquele que, significativamente, apresentar maior probabilidade de sobrevivência dos pacientes.

REFERÊNCIAS

- ALLISON, PD. Event history analysis. **London: Sage University Paper 46**, 1984.
- ARMITAGE, P, BERRY, G. Statistical methods in medical research. **USA:3 ed. Blackwell Science Ltda.**, 1994.
- BAILAR III, JC, MOSTELLER, F. **Medical uses of statistics**. Boston: 2 ed. NEJM Book, 1992.
- CHIESA, EMF. **Análise de Sobrevivência através de um método não paramétrico** [monography]. Santa Maria: Curso de Especialização em Estatística e Modelagem Quantitativa /UFSM; 2002.
- COLOSIMO, EA. Análise de Sobrevivência Aplicada. In: **Anais 9º, SEAGRO, REUNIÃO ANUAL DA RBRAS, 46ª, Piracicaba, SP**. Brasil. 2001.
- COX, DR. Regression models and life tables. **Journal of the Royal Statistical Society B**, 1972; v. 34, p.187-220.
- COX, DR; OAKES, D. Analysis of survival data. **London: Chapman and Hall**, 1984.
- ECHEVESTE, SS. **Análise de sobrevivência: um estudo na área educacional** [monography]. Porto Alegre: Instituto de Matemática, Departamento de Estatística/ UFRGS;1997.
- ELANDT-JOHNSON, RC, JOHNSON, NL. Survival models and data analysis. **New York: John Wiley and Sons**, 1980.
- FISCHER, LD., BELLE, GV. Biostatistics – A methodology for the health sciences. **New York: John Wiley & Sons, Inc.**, 1995.
- HOSMER, DW Jr, LEMESHOW, S. Applied survival analysis – Regression modeling of time to event data. **New York: John Wiley & Sons, Inc**, 1998.
- KALBFLEISCH, JD, PRENTICE, RL. The statistical analysis of failure time data. **New York: John Wiley and Sons**, 1980.
- KAPLAN, EL., MEIER, P. Nonparametric estimation for incomplete observations. **Journal of the American Statistical Association**, 1958; v. 53, n.282, p.457-481.
- KLEINBAUM, DG. Survival analysis – A self learning text. **New York: Verlag New York, Inc**, 1996.
- LAWLESS, JF. Statistical models and methods for lifetime data. **New York: John Wiley and Sons**, 1982.

LE, CHAP T., Applied survival analysis. **John Wiley & Sons, Inc.** New York, 1997.

Excel. **Microsoft Corporation.** Quito, versão 7, 2001.

NELSON, W., Accelerated life testing: statistical models, data analysis and test plans. **New York: John Wiley and Sons**, 1990.

RODRIGUES, PC. **Bioestatística.** Niterói: 2-ed EDUFF, 1993.

SOARES, JF., COLOSIMO, EA. Métodos estatísticos na pesquisa clínica. In: **Anais 6º, SEAGRO, REUNIÃO ANUAL DA RBRAS, 40º, Ribeirão Preto, SP.** Brasil.1995.

STATISTICA for windows. **Statsoft.** Tulsa, 1998, versão 5.1.