Universidade de São Paulo
Faculdade de Saúde Pública

**Avaliação do efeito das perdas de seguimento nas
análises feitas pelo estimador produto-limite de
Kaplan-Meier e pelo modelo de riscos proporcionais
de Cox**

Marcia Moreira Holcman

Tese apresentada ao Programa de Pós-Graduação em Saúde Pública para obtenção do título de Doutor em Saúde Pública.

Área de Concentração: Epidemiologia
Orientadora: Profa Dra Maria do
Rosário Dias de Oliveira Latorre

São Paulo

2006

**Avaliação do efeito das perdas de seguimento nas
análises feitas pelo estimador produto-limite de
Kaplan-Meier e pelo modelo de riscos proporcionais
de Cox**

Marcia Moreira Holcman

Tese apresentada ao Programa de Pós-Graduação em Saúde Pública da Faculdade de Saúde Pública da Universidade de São Paulo para obtenção do título de Doutor em Saúde Pública.

Área de Concentração: Epidemiologia
Orientadora: Profa Dra Maria do Rosário Dias de Oliveira Latorre

São Paulo

2006

AGRADECIMENTOS

Agradeço à Professora. Dra. Maria do Rosário Dias de Oliveira Latorre pela orientação.

Aos funcionários do Serviço de Arquivo Médico – SAME e do Registro de Câncer – RHC do Hospital do Câncer de São Paulo pela atenção e auxílio na pesquisa dos prontuários médicos.

RESUMO

Holcman MM. **Avaliação do efeito das perdas de seguimento nas análises feitas pelo estimador produto-limite de Kaplan-Meier e pelo modelo de riscos proporcionais de Cox.** São Paulo;2006. [Tese de Doutorado -Faculdade de Saúde Pública da USP]

Introdução: As técnicas mais comumente empregadas em análise de sobrevida que utilizam dados censurados são o estimador produto limite de Kaplan-Meier (KM) e o modelo de riscos proporcionais de Cox. Estas técnicas têm como suposição que a causa da perda de seguimento seja independente do tempo de sobrevida. **Objetivo:** O presente estudo visa a analisar o efeito das perdas de seguimento nestas duas técnicas. **Material e Métodos:** O estudo foi realizado utilizando-se o banco de dados dos pacientes cadastrados no Registro Hospitalar do Hospital do Câncer de São Paulo em 1994. Foram elaborados 28 bancos de dados simulando perdas informativas e não informativas. A perda informativa foi simulada transformando os óbitos em vivos, na proporção de 5 a 50%. A perda não informativa foi simulada através do sorteio de 5 a 50% do total do banco. O estimador de Kaplan-Meier (KM) foi utilizado para estimar a sobrevida acumulada no primeiro, terceiro e quinto ano de seguimento, e o modelo de riscos proporcionais de Cox para estimar as *hazard ratio* (HR). Todas as estimativas obtidas no KM e as HR's foram comparadas com os resultados do banco de dados original. **Resultados:** Houve maior proporção de perda nos pacientes com maior escolaridade, admitidos por convênio e particular e os menos graves (estádio I ou II). Quanto maior a proporção de perda informativa, maior a diferença alcançada nas estimativas realizadas pelo KM, verificando-se que a perda de seguimento superior a 15% acarretou diferenças superiores a 20% nas estimativas da probabilidade de sobrevida. As HR's foram menos afetadas, e proporções superiores a 20% de perda de seguimento acarretaram variações de cerca de 10% nas estimativas. Quando as perdas foram não informativas não houve diferenças

significativas nas estimativas pelo KM e nas HR's em relação ao banco original. **Conclusões:** É importante avaliar se as perdas ocorridas em estudos de coorte são informativas ou não, pois se forem podem acarretar distorções principalmente nas estimativas feitas pelo método de KM.

Descritores: Análise de Sobrevida, Kaplan-Meier, modelo de Cox, perda de seguimento

SUMMARY

Holcman MM.: **The impact of the loss to follow-up when using the Kaplan Meier estimator and the Cox proportional hazard model.** [Tese de Doutorado - Faculdade de Saúde Pública da Universidade de São Paulo]

Introduction: The Kaplan Meier product limit estimator (KM) and the Cox proportional hazard (HR) model are the most used tools in survival analysis. These two methods have the key assumption that censoring must be independent from the survival time. **Objective:** To analyze the consequences of loss to follow up in these two methods. **Methods:** The study has utilized the data of the Cancer Registry of the patients of Hospital do Cancer in São Paulo of 1994. The informative censure was simulated transforming the death by 5 to 50% into alive. Besides 5 to 50% was spared at random simulating the non-informative censoring. The survival probability and was calculated to the first, third and fifth year of follow –up. All the estimated probabilities and HR's were compared with the results of the original data. **Results:** Patients with greater scholars, lower stages and admitted by health plans or private had more losses to follow up. The maximum proportion of accepted loss to follow –up is 10% to 15% when using the KM estimator, and the HR are less affected by the loss to follow-up and one can afford having 20% of it. When the losses were non informative there were no differences between the original probabilities. **Conclusions:** The possibility of over or under estimated probability must be analyzed in the presence of the losses to follow- up when using the KM and HR in survival analyses.

ÍNDICE

1	INTRODUÇÃO	1
1.1	As Censuras ou tempos incompletos	3
2	OBJETIVOS	8
3	MATERIAL E MÉTODOS	9
3.1	Local do Estudo e casuística	9
3.2	Bases de dados	9
3.2.1	Elaboração das bases de dados simuladas	10
3.2.1.1	Simulação da perda informativa nos óbitos	10
3.2.1.2	Simulação da perda informativa nos vivos	11
3.2.1.3	Simulação da perda não informativa	11
3.2.1.4	Simulação sem perdas ao longo do seguimento	11
3.3	Variáveis de estudo	11
3.4	Análise estatística	12
3.5	Questões Éticas	13
4	RESULTADOS	14
4.1	Revisão da base de dados original	14
4.2	Simulação das bases de dados	15
4.2.1	Perdas informativas nos óbitos	15

4.2.2 Perdas informativas nos vivos	15
4.2.3 Perdas não informativas	16
4.2.4 Sem perdas anuais de 1994 a 1999	16
4.3 Análise das perdas	18
4.4 Caracterização da amostra	21
4.5 Análise da probabilidade de sobrevida acumulada e HR no total dos pacientes	22
4.6 Análise do efeito das perdas nas probabilidades de sobrevida acumuladas nas bases de dados	24
4.6.1 Total de pacientes	25
4.6.2 Variável sexo	19
4.6.3 Variável Idade	33
4.6.4 Variável Cor/raça	38
4.6.5 Variável Escolaridade	42
4.6.6 Variável Categoria de admissão	46
4.6.7 Variável Estádio	50
4.7 Análise do efeito das perdas nas HR's	54
4.8 Análise do efeito das perdas dos anos de 1994 a 1999 na probabilidade acumulada de sobrevida no total dos pacientes	62
4.9 Análise do efeito das transformações das perdas de seguimento em vivo ao final do estudo na probabilidade acumulada de sobrevida e nas HR's no total dos pacientes	71
5 DISCUSSÃO	74

6 CONCLUSÕES	84
7 REFERENCIAIS BIBLIOGRÁFICAS	85
ANEXOS	87

Anexo 1 Ficha do Registro de Câncer de 1994

Anexo 2 Aprovação da comissão de ética Hospital do Câncer

Anexo 3 Aprovação da comissão de ética da Faculdade de Saúde Pública

Anexo 4 Tabela das probabilidades de sobrevida acumulada do total de pacientes com perdas informativas nos óbitos até o primeiro, terceiro e quinto ano de seguimento, RHC, 1994

Anexo 5 Tabela das probabilidades de sobrevida acumulada do total de pacientes com perdas não informativas até o primeiro, terceiro e quinto ano de seguimento, RHC, 1994

Anexo 6 Tabela das probabilidades de sobrevida acumulada do total de pacientes com perdas informativas nos óbitos segundo sexo até o primeiro, terceiro e quinto ano de seguimento, RHC, 1994

Anexo 7 Tabela das probabilidades de sobrevida acumulada segundo sexo com perdas não informativas até o primeiro, terceiro e quinto ano de seguimento, RHC, 1994

Anexo 8 Tabela das probabilidades de sobrevida acumulada segundo idade com perdas informativas nos óbitos até o primeiro, terceiro e quinto ano de seguimento, RHC, 1994

Anexo 9 Tabela das probabilidades de sobrevida acumulada segundo idade com perdas não informativas até o primeiro, terceiro e quinto ano de seguimento, RHC, 1994

Anexo 10 Tabela das probabilidades de sobrevida acumulada segundo cor/raça com perdas informativas nos óbitos até o primeiro, terceiro e quinto ano de seguimento, RHC, 1994

Anexo 11 Tabela das probabilidades de sobrevida acumulada segundo cor/raça com perdas não informativas até o primeiro, terceiro e quinto ano de seguimento, RHC, 1994

Anexo 12 Tabela das probabilidades de sobrevida acumulada segundo

escolaridade com perdas informativas nos óbitos até o primeiro, terceiro e quinto ano de seguimento, RHC, 1994

Anexo 13 Tabela das probabilidades de sobrevida acumulada segundo escolaridade com perdas não informativas até o primeiro, terceiro e quinto ano de seguimento, RHC, 1994

Anexo 14 Tabela das probabilidades de sobrevida acumulada segundo categoria com perdas informativas nos óbitos até o primeiro, terceiro e quinto ano de seguimento, RHC, 1994

Anexo 15 Tabela das probabilidades de sobrevida acumulada segundo categoria com perdas não informativas até o primeiro, terceiro e quinto ano de seguimento, RHC, 1994

Anexo 16 Tabela das probabilidades de sobrevida acumulada segundo estadio com perdas informativas nos óbitos até o primeiro, terceiro e quinto ano de seguimento, RHC, 1994

Anexo 17 Tabela das probabilidades de sobrevida acumulada segundo estadio com perdas não informativas até o primeiro, terceiro e quinto ano de seguimento, RHC, 1994

INDICE DE TABELAS

Tabela 1 - Número e porcentagem de pacientes segundo status de seguimento inicial e final, RHC, 1994.....	14
Tabela 2 – Nome, descrição geral e número de observações das bases de dados geradas.....	17
Tabela 3 – Número e porcentagem de pacientes segundo status final e características demográficas e clínicas, RHC, 1994.....	19
Tabela 4 Número e porcentagem dos pacientes com status inicial de perdidos de seguimento segundo status final e ano de seguimento, RHC, 1994.....	20
Tabela 5 -Número e percentual de pacientes segundo recuperação de informação no prontuário e características demográficas e clínicas , RHC, 1994.....	21
Tabela 6 – Número de distribuição percentual dos pacientes segundo características demográficas e clínicas e recuperação de seguimento, RHC, 1994.....	22
Tabela 7 – Probabilidade de sobrevida acumulada do total de pacientes segundo as características demográficas e clínicas até o primeiro, terceiro e quinto ano de seguimento, RHC, 1994.....	23
Tabela 8 – Razão das funções de riscos segundo as características demográficas e clínicas, RHC, 1994.....	24
Tabela 9 – Modelos de regressão linear simples segundo as perdas de 5 a 50% e variação percentual das probabilidades de sobrevida acumuladas no total de pacientes, RHC,1994.....	26
Tabela 10 - Modelos da regressão linear simples segundo as perdas de 5 a 50% e variação percentual das probabilidades de sobrevida acumuladas até o primeiro, terceiro e quinto ano de seguimento segundo sexo, RHC,1994.....	30
Tabela 11 – Modelos de regressão linear simples segundo as perdas de 5 a 50% e variação percentual das probabilidades de sobrevida acumuladas até o primeiro, terceiro e quinto ano de seguimento segundo idade, RHC,1994.....	35
Tabela 12 - Modelos de regressão linear simples segundo as perdas de 5 a 50% e variação percentual das probabilidades de sobrevida acumuladas até o primeiro, terceiro e quinto ano de seguimento segundo cor/raça, RHC,1994.....	39
Tabela 13 – Modelos de regressão linear simples segundo as perdas de 5 a 50% e variação percentual das probabilidades de sobrevida acumuladas até o primeiro, terceiro e quinto ano de seguimento segundo escolaridade, RHC,1994.....	43

Tabela 14 Modelos de regressão linear simples segundo as perdas de 5 a 50% e variação percentual das probabilidades de sobrevida acumuladas até o primeiro, terceiro e quinto ano de seguimento segundo categoria de admissão, RHC,1994.....	47
Tabela 15 – Modelos de regressão linear simples segundo as perdas de 5 a 50% e variação percentual das probabilidades de sobrevida acumuladas até o primeiro, terceiro e quinto ano de seguimento segundo estadio clínico, RHC,1994.....	51
Tabela 16 – Modelos regressão linear simples segundo as perdas de 5 a 50% e variação percentual razões das funções de risco segundo as variáveis de estudo, RHC,1994.....	55
Tabela 17 Modelos de regressão linear simples segundo os bancos sem perdas de 1994 a 1999 e variação percentual das probabilidades de sobrevida acumuladas segundo idade, RHC, 1994.....	63
Tabela 18 – Probabilidade de sobrevida acumulada do total de pacientes segundo as características demográficas e clínicas e variação percentual quando os perdidos são considerados vivos ao final de seguimento, RHC, 1994.....	72
Tabela 19 – Razão das funções de risco do total de pacientes segundo as características demográficas e clínicas e variação percentual quando os perdidos são considerados vivos ao final de seguimento, RHC, 1994.....	73

ÍNDICE DE FIGURAS

Figura 1 – Esquema de seguimento e análise em estudos de sobrevida.....	2
Figura 2 – Gráficos das probabilidades de sobrevida acumulada segundo o total de pacientes até o primeiro, terceiro e quinto ano de seguimento nos bancos gerados com perdas , RHC, 1994...	27
Figura 3 – Estimativas das probabilidades de sobrevida para 4 bases de dados	28
Figura 4 – Gráficos das probabilidades de sobrevida acumulada segundo o sexo até o primeiro, terceiro e quinto ano de seguimento nos bancos gerados com perdas , RHC, 1994.....	31
Figura 5 – Estimativas das probabilidades de sobrevida para 4 bases de dados	32
Figura 6 – Gráficos das probabilidades de sobrevida acumulada segundo a idade até o primeiro, terceiro e quinto ano de seguimento nos bancos gerados com perdas, RHC, 1994.....	36
Figura 7 – Estimativas das probabilidades de sobrevida para 4 bases de dados	37
Figura 8 – Gráficos das probabilidades de sobrevida acumulada segundo cor/raça até o primeiro, terceiro e quinto ano de seguimento nos bancos gerados com perdas, RHC, 1994.....	40
Figura 9 – Estimativas das probabilidades de sobrevida para 4 bases de dados	41
Figura 10 – Gráficos das probabilidades de sobrevida acumulada segundo a escolaridade até o primeiro, terceiro e quinto ano de seguimento nos bancos gerados com perdas, RHC, 1994.....	44
Figura 11 – Estimativas das probabilidades de sobrevida para 4 bases de dados	45
Figura 12 – Gráficos das probabilidades de sobrevida acumulada segundo a categoria até o primeiro, terceiro e quinto ano de seguimento nos bancos gerados com perdas, RHC 1994.....	48
Figura 13 – Estimativas das probabilidades de sobrevida para 4 bases de dados	49
Figura 14 – Gráficos das probabilidades de sobrevida acumulada segundo estadio até o primeiro, terceiro e quinto ano de seguimento nos bancos gerados com perdas, RHC, 1994.....	52
Figura 15 – Estimativas das probabilidades de sobrevida para 4 bases de dados	53
Figura 16 – Razão das funções de riscos segundo sexo, RHC , 1994.....	56

Figura 17 – Razão das funções de riscos segundo idade, RHC, 1994.....	57
Figura 18 – Razão das funções de riscos segundo cor/raça,RHC, 1994.....	58
Figura 19 – Razão das funções de riscos segundo escolaridade, RHC, 1994.....	59
Figura 20 – Razão das funções de riscos segundo categoria, RHC , 1994.	60
Figura 21 – Razão das funções de riscos segundo estágio, RHC , 1994.	61
Figura 22 – Gráficos das probabilidades de sobrevida acumulada segundo sexo até o primeiro, terceiro e quinto ano de seguimento nos bancos gerados sem perdas de 1994 a 1999, RHC, 1994.....	64
Figura 23 – Gráficos das probabilidades de sobrevida acumulada segundo idade até o primeiro, terceiro e quinto ano de seguimento nos bancos gerados sem perdas de 1994 a 1999, RHC, 1994.....	65
Figura 24 – Gráficos das probabilidades de sobrevida acumulada segundo raça até o primeiro, terceiro e quinto ano de seguimento nos bancos gerados sem perdas de 1994 a 1999, RHC , 1994.....	66
Figura 25 – Gráficos das probabilidades de sobrevida acumulada segundo escolaridade até o primeiro, terceiro e quinto ano de seguimento nos bancos gerados sem perdas de 1994 a 1999, RHC, 1994.....	67
Figura 26 – Gráficos das probabilidades de sobrevida acumulada segundo categoria até o primeiro, terceiro e quinto ano de seguimento nos bancos gerados sem perdas de 1994 a 1999, RHC, 1994.....	68
Figura 27 – Gráficos das probabilidades de sobrevida acumulada segundo estadio até o primeiro, terceiro e quinto ano de seguimento nos bancos gerados sem perdas de 1994 a 1999, RHC, 1994.....	69
Figura 28 – Gráficos das probabilidades de sobrevida acumulada do total da base de dados até o primeiro, terceiro e quinto ano de seguimento nos bancos gerados sem perdas de 1994 a 1999, RHC, 1994.....	70

1 INTRODUÇÃO

Em saúde, nos estudos de coorte, sejam eles observacionais ou ensaios clínicos, a medida do tempo até a ocorrência de um evento é uma das variáveis de interesse. As técnicas de análise de sobrevida são empregadas para se calcular as probabilidades de sobrevida em diversos momentos de seguimento, bem como o efeito dos fatores associados à ocorrência do evento de interesse ao longo do tempo (BUSTAMANTE e col. 2002; LEE e col. 1992).

A sobrevida é a medida do tempo (que pode ser medido em anos, meses ou semanas) desde a entrada do indivíduo no estudo até a ocorrência do evento de interesse (falha).

A entrada do indivíduo pode se dar, por exemplo, pelo diagnóstico de uma doença, realização de uma cirurgia, início de um tratamento ou nascimento. Os indivíduos podem entrar em diferentes tempos-calendário e a sobrevida é considerada a partir da entrada no estudo (tempo zero). A falha é qualquer evento de interesse como a morte, a ocorrência de um efeito adverso a um medicamento ou a remissão da doença.

A principal característica da análise de sobrevida é a de que alguns indivíduos possam permanecer sem falha ao término do estudo, possam sofrer outro evento (como morte por outra causa que não a de interesse), possam abandonar o estudo ou serem perdidos de seguimento. Neste caso, o tempo de acompanhamento é interrompido e estas observações são censuradas à direita, pois o tempo não é medido até a ocorrência do desfecho.

Os inícios são truncados à esquerda, isto é, o tempo de observação de cada indivíduo não leva em conta o que aconteceu no passado (LATORRE 1996; BUSTAMANTE e col. 2002; LATORRE 2005). A Figura 1 representa o esquema de acompanhamento de um grupo de 4 indivíduos que são acompanhados de janeiro de 1990 até março de 1991.

Figura 1 – Esquema de seguimento e análise em estudos de sobrevida.

Neste esquema, os indivíduos marcados com **X** sofreram o evento de interesse (falha) e os marcados com **#** não sofreram a falha, logo são considerados censurados. Na primeira parte da Figura 1, pode-se verificar que o indivíduo 1 entrou em fevereiro de 1990 (tempo 0) e sofreu a falha em março de 1991 e, assim, o tempo de sobrevida dele é 12 meses; já o indivíduo 2 entrou no estudo em janeiro de 1990 e não sofreu falha, sendo, portanto, censurado em outubro de 1992. O indivíduo 3 entrou em março de 1990 e teve óbito por outra causa em fevereiro de 1991, sendo, então, acompanhado por 10 meses. O indivíduo 4 entrou em março de 1990 e sofreu o evento em dezembro de 1990, totalizando o tempo de sobrevida de 6 meses (HOSMER e LEMESHOW 1999).

1.1 As censuras ou tempos incompletos

Os mecanismos que levam aos tempos de seguimento incompletos ou censuras podem ocorrer antes do início da observação. Por exemplo, quando se estuda a sobrevida de pacientes infectados pelos vírus das hepatites B ou C ou com o vírus HIV, onde o momento exato da contaminação por estes vírus não pode ser obtido através dos testes diagnósticos atuais. Estes casos são denominados censura à esquerda (KLEINBAUM 1995).

Algumas vezes, as falhas podem ocorrer em um intervalo de tempo quando os pacientes são visitados em períodos pré-determinados, e neste caso, são denominados censura no intervalo. LINDSEY e RYAN (1998) citam que quando os dados são censurados no intervalo, é comum assumir que a falha ocorra no meio, no início ou ao final do intervalo entre as visitas. Uma vez estimados estes tempos até a falha, os pesquisadores costumam utilizar os métodos usuais de análise de sobrevida. Segundo os autores, este procedimento pode acarretar distorções nas estimativas geradas e recomendam o uso de métodos paramétricos na análise de sobrevida em dados com este tipo de censura.

Na maior parte dos estudos, os tempos incompletos aparecem no final do período de observação, nestes casos, a falta de acompanhamento até a falha se situa à direita e são chamados censura à direita. Estes tipos de censura ocorrem quando:

- a) Os indivíduos terminam o estudo sem apresentar a falha, ou seja, em estudos de sobrevida onde pacientes com câncer são seguidos por 5 anos a partir do diagnóstico e, ao final deste período, uma parte deles ainda continua viva (a falha neste caso é o óbito por câncer);
- b) No caso anterior pode acontecer de alguns pacientes serem perdidos de seguimento, por exemplo, quando o paciente decide parar o tratamento por algum efeito adverso ou se muda de cidade e decide continuar o tratamento em outro local, tornando difícil sua localização para identificar sua situação de vivo ou óbito ao final do período de seguimento;
- c) O paciente do estudo pode falecer de outra doença que não o câncer e ter seu tempo de acompanhamento encerrado.

HOUGAARD (1999) em seu artigo sobre fundamentos da análise de sobrevida, comenta que a presença de censura constitui um dos grandes problemas técnicos da análise de sobrevida e salienta que a censura não pode ser relacionada com o tempo de sobrevida, isto é, o resíduo da distribuição da probabilidade de sobrevida para os indivíduos censurados tem que ser igual para os indivíduos não censurados.

ALTMAN e col. (1995) realizaram uma revisão sistemática dos trabalhos que utilizam as técnicas de análise de sobrevida publicados em 5 periódicos de oncologia clínica, no período de outubro a dezembro de 1991, e detectaram diversos problemas envolvendo alguns aspectos das análises, tais como, não informação do tamanho da amostra, tempo de seguimento não definido, falta de clareza com relação às variáveis independentes utilizadas no estudo, informações sobre dados incompletos e outros. Com

relação ao tempo de seguimento, poucos artigos apresentam uma definição exata do início e término do estudo. Em especial, destacam que, na maior parte dos trabalhos analisados, não há informação clara sobre o percentual de perdas ao longo do estudo e como elas foram tratadas.

MOHER e col. (2001) realizaram a revisão do CONSORT (*Consolidated Standards of Reporting Trials*) que abrange as recomendações para a melhoria da qualidade nas informações contidas nos artigos relativos a ensaios clínicos aleatorizados. Dentre elas, está a sugestão de informar o número de pacientes, incluindo o número de perdas, visando a aumentar a confiabilidade destas comunicações, uma vez que, a falta delas estaria associada a vícios na estimação dos efeitos do tratamento.

LEUNG e col. (1997) em artigo de revisão, analisaram todas as possibilidades da ocorrência de tempos de observação incompletos e constataram que, na prática, existem algumas estratégias para se trabalhar com dados que envolvem observações incompletas (censuras). Uma delas seria ignorar esses tempos incompletos e utilizar somente casos com observação completa. Esta decisão, apesar de simples, ocasionaria várias conseqüências, dentre elas a perda de eficiência pela diminuição do tamanho da amostra, pois em muitos estudos epidemiológicos é comum a ocorrência de 50% ou mais observações censuradas. e o viés nas estimativas baseadas somente nas observações não censuradas.

Outra alternativa apontada seria completar os dados censurados por meio de imputação, e no contexto da censura à direita, existiriam duas maneiras para se completar os tempos; uma supondo que a falha ocorra logo após a perda, ou utilizando-se um modelo para estimar o tempo residual até a falha. A última técnica citada para se analisar dados censurados seria utilizar métodos de máxima verossimilhança que levassem em conta os mecanismos de censura. Estas técnicas, que incluem o estimador de Kaplan-Meier, e a regressão de Cox, são consideradas pelo autor as mais eficazes por utilizarem toda a informação disponível. Porém,

também com estas técnicas são necessárias algumas suposições sobre o mecanismo de censura para sua utilização.

CLARK e col. (2002) propuseram uma medida para a quantificação da completude do tempo de seguimento, que pode ser utilizada tanto em ensaios clínicos como em estudos de coorte. Com ela, ao invés da simples contagem das perdas, que não leva em conta o tempo, teríamos uma idéia aproximada do percentual perdido do tempo, devido às censuras. Esta medida seria a razão entre a soma do total de pessoas-tempo e o total potencial de tempo do estudo (soma do tempo de entrada no estudo até seu término ou ser for o caso até a ocorrência do evento) em percentual. Quanto maior o seu valor mais completo seria o estudo.

Os autores sugerem, também, melhorar a qualidade das informações relativas as análises de estudos de sobrevida e identificar sub-grupos com problemas de seguimento, comparando estas medidas, com relação aos vários braços do estudo (no caso de ensaios clínicos) ou com relação às exposições (no caso de estudos de coorte).

LAGAKOS (1979), em seu artigo sobre censura a direita, afirma que a maior parte dos métodos de análise de sobrevida tem como suposição básica a censura não informativa. O exemplo mais simples, deste tipo de censura, seria o ensaio clínico onde uma parte dos pacientes apresenta o desfecho e o restante termina o estudo sem apresentar a falha. Outros casos, em que alguns pacientes saem do estudo por razões relacionadas ao tipo de terapia e tem seu acompanhamento terminado, podem ser consideradas censuras do tipo informativas, pois a razão da censura, poderia estar relacionada com o tempo de sobrevida futuro. Ao final do artigo, o autor declara que, quando o percentual de censura informativa é pequeno, as estimativas baseadas em métodos para censura não informativa sofrem pequeno desvio mas se esta proporção for grande é necessário a utilização de outros métodos para censura informativa.

Os estudos com censura informativa devem ser analisados através de outras técnicas, visto que, segundo LINK (1989), as técnicas de Kaplan-Meier podem superestimar as probabilidades de sobrevida, caso o motivo da censura indique um prognóstico desfavorável para a sobrevida. O modelo proposto pelo autor, deve ser utilizado quando a censura ocorre em subpopulações com alto risco e é definida por uma função de fragilidade.

LAGAKOS e col. (1978) propõem um modelo com um parâmetro escalar θ (variando de zero até um) que refletiria o quanto a sobrevida seria afetada pela censura. E, neste modelo, para valores de $\theta=1$, a função de verossimilhança não é afetada pela censura, e para valores de θ mais próximos de 0, os efeitos das censuras seriam mais importantes nas estimativas da função de sobrevida.

Na revisão bibliográfica realizada pela autora desta tese, não foram encontrados artigos que analisem especificamente a questão do percentual de perdas de seguimento. SIMON e WITTES (1986) comentam em seu editorial sobre recomendações para se publicar ensaios clínicos que, de uma maneira arbitrária, pode-se aceitar até 15% de perdas de seguimento e que perdas iguais ou superiores que 20% refletem problemas na seleção dos pacientes. No entanto, na maioria dos estudos, não há referências de como foram tratadas as perdas de seguimento, não incorporando informações suficientes para a real avaliação do impacto destas perdas na estimativa do tempo de sobrevida. E foi este fato que inspirou a realização deste trabalho.

2 OBJETIVOS

2.1 Geral

Analisar o efeito das perdas de seguimento nas probabilidades de sobrevida acumulada calculadas pelo estimador produto-limite de Kaplan-Meier e nas razões das funções de risco (*hazard ratio*) estimadas pelo modelo de riscos proporcionais de Cox.

2.2 Específicos

- ❑ Descrever a elaboração dos bancos de dados das simulações;
- ❑ Analisar as proporções das perdas de seguimento com relação às variáveis: sexo, idade, cor/raça, escolaridade, categoria de admissão e estágio clínico;
- ❑ Analisar o efeito da perda de seguimento do tipo informativa nos estimadores de KM e nas HR's, segundo as variáveis: sexo, idade, cor/raça, escolaridade, categoria de admissão e estágio clínico;
- ❑ Analisar o efeito da perda de seguimento do tipo não informativa nos estimadores de KM e nas HR's, segundo as variáveis: sexo, idade, cor/raça, escolaridade, categoria de admissão e estágio clínico.

3. MATERIAL E MÉTODOS

3.1. Local do Estudo e casuística

O estudo foi realizado utilizando o banco de dados dos pacientes do Registro Hospitalar de Câncer (RHC) do Hospital A. C. Camargo, da Fundação Antonio Prudente, em São Paulo, referente ao ano de 1994.

Foram analisados todos os pacientes com diagnóstico de câncer registrados no período de 01/01/1994 a 31/12/1994, com base nas informações coletadas pelo próprio RHC (Anexo 1).

3.2. Bases de Dados

A base de dados fornecida pelo RHC continha informações de 2842 pacientes. Foram excluídos 558 pacientes que tinham data de diagnóstico diferente de 1994, resultando 2284 pacientes.

O primeiro passo foi atualizar a data da última observação dos pacientes codificados como vivos (códigos de 1 a 3) ou perdidos de seguimento (código 6). A revisão destas informações foi realizada nos prontuários médicos, no período de março de 2003 a maio de 2004.

A falha de interesse foi o óbito e o período de seguimento definido em cinco anos. Os pacientes tiveram seu tempo de sobrevida estimado da data do diagnóstico até a ocorrência do óbito ou até 60 meses. Deste modo, foram considerados censurados todos os pacientes que continuavam vivos após 5 anos de seguimento, e como perdidos, todos os pacientes que não foram localizados pelo serviço neste período de seguimento.

Após todos os procedimentos de avaliação e atualização das informações nos prontuários médicos, chegou-se ao banco de dados que seria utilizado em qualquer análise de sobrevida. Esta base de dados foi denominada, neste trabalho, como **base de dados padrão** e foi a base

inicial para a avaliação do efeito das perdas de seguimento em estudos de sobrevida.

O detalhamento desta revisão consta no item 4.1 dos resultados (páginas 15 e 16).

3.2.1 Elaboração das bases de dados simuladas

Foram elaboradas 28 bases de dados para este trabalho. A base de dados padrão foi separada em 3 sub-amostras: uma com os pacientes considerados vivos no período de cinco anos (exceto os perdidos de seguimento), a segunda com os perdidos de seguimento e a terceira com todos os pacientes que foram a óbito no período de cinco anos. A descrição da elaboração das bases de dados está a seguir e a apresentação das mesmas está presente no item 4.2 dos resultados (páginas 16 a 20).

3.2.1.1 Simulação da perda informativa nos óbitos

A partir da base de dados dos óbitos foram geradas 10 sub-amostras com 5%, 10%, 15%, 20%, 25%, 30%, 35%, 40%, 45% e 50% das observações. Estes pacientes sorteados foram recodificados como censuras. A junção de cada uma destas sub-amostras com a base de dados dos vivos formou outras 10 bases de dados, cada uma delas contendo o total de vivos, mais $x\%$ (sendo x de 5 a 50) de óbitos transformados em censuras, mais $(100-x)\%$ de óbitos.

A construção deste grupo de 10 bases de dados teve como objetivo simular as perdas de seguimento informativas, pois foram recodificados como censura pacientes que sabidamente foram a óbito.

3.2.1.2 Simulação da perda informativa nos vivos

Outra base de dados foi construída transformando os pacientes perdidos de seguimento como vivos após 60 meses. Ao assumir como vivos, o tempo de seguimento de todos estes pacientes foi modificado para 5 anos e estes pacientes foram considerados censura. Este procedimento simulou outro tipo de perda de seguimento informativa, pois todos foram considerados como vivos.

3.2.1.3 Simulação da perda não informativa

Outro grupo de 10 sub-amostras foi criado a partir de sorteio nas proporções de 5%, 10%, 15%, 20%, 25%, 30%, 35%, 40%, 45% e 50% do total da base de dados padrão, mas desta vez, os pacientes sorteados foram excluídos da base de dados padrão. Este procedimento visou a construir as 10 bases de dados que foram consideradas como tendo perdas não informativas ao longo do seguimento.

3.2.1.4 Bases sem perdas ao longo do seguimento

Por fim, outro grupo de 6 bases de dados foi construído, excluindo-se as perdas de seguimento ocorridas em cada ano de seguimento (1994 a 1999).

3.3 Variáveis de estudo

Foram selecionadas as seguintes variáveis da base de dados e que constam da ficha de registro (Anexo 1). Estas variáveis foram selecionadas por serem as mais utilizadas em estudos de sobrevida de pacientes com câncer:

1. Sexo;

2. Idade: analisada em quatro categorias (0 a 14 anos, 15 a 39 anos, 40 a 64 anos e 65 anos ou mais);
3. Cor/ Raça: analisada em duas categorias (branca ou amarela e parda ou preta);
4. Escolaridade: analisada em duas categorias (até fundamental e ensino médio ou superior);
5. Categoria de admissão: analisada em duas categorias (SUS e convênio ou particular);
6. Estádio Clínico: analisado em duas categorias (I + II e III + IV).

3.4 Análise estatística

O teste de associação pelo qui-quadrado foi utilizado para avaliar a associação entre as variáveis de estudo e o status final do paciente como perdido ou não de seguimento.

Na análise de sobrevida, utilizou-se a técnica de Kaplan-Meier (KM) e foram apresentadas as estimativas das probabilidades de sobrevida acumuladas até o primeiro, terceiro e quinto ano de seguimento. As curvas das probabilidades de sobrevida foram comparadas utilizando-se o teste *log-rank*. Para avaliação visual das diferenças nas probabilidades de sobrevida estimadas, segundo os tipos de perdas simuladas, foram construídas 4 curvas de sobrevida de Kaplan-Meier :

1. para a base de dados padrão,
2. para a base de dados padrão sem perdas de seguimento;
3. para a base de dados com 50% de perdas informativas nos óbitos ;
4. para a base de dados com 50% de perdas não informativas.

O modelo de regressão de Cox foi utilizado para se estimar as razões das funções de risco (HR) e os respectivos intervalos de 95% de confiança para cada uma das variáveis de estudo. Em todos os modelos foram utilizadas como categorias de referência aquelas que apresentavam menor força de mortalidade.

Para se analisar as variações sofridas nas probabilidades de sobrevida acumuladas (no 1º, 3º e 5º ano de seguimento) e nas HR's, foram construídos diagramas lineares, um para cada tipo de perda.

Para se estimar a tendência de queda ou acréscimo nos valores obtidos nas estimativas feitas nas análises segundo KM e nas HR's, foram realizadas análises de regressão linear simples para cada tipo de perda e cada uma das variáveis de estudo. O valor da probabilidade de sobrevida acumulada e a estimativa pontual das HR's foram as variáveis dependentes e a variável independente foi a proporção de perda gerada.

Foram calculadas também as variações percentuais através da fórmula:

$$\text{Variação Percentual (de 0 a } x\%) = \left(\frac{\text{probabilidade com perda de } x\%}{\text{probabilidade sem a perda}} - 1 \right) * 100$$

Em todas as análises utilizou-se o nível de significância de 5%.

3.5 Questões éticas

Este é um estudo retrospectivo, em que os pesquisadores se comprometeram a manter sigilo sobre a identificação do paciente. O projeto foi aprovado pelas comissões de ética da Faculdade de Saúde Pública e do Hospital do Câncer (Anexos 2 e 3).

4. RESULTADOS

4.1 Revisão da base de dados original

Na base de dados inicial (total de 2284 pacientes) havia 1117 codificados como vivos, 536 codificados como perdidos de seguimento e 631 codificados como óbito. Após a revisão dos prontuários, dos 1117 pacientes classificados como vivos, 661 ainda permaneciam vivos (conservando a codificação original), 201 foram a óbito (recodificados para óbito) e 255 foram perdidos de seguimento (não compareceram para consulta de retorno após 3 meses da data agendada no prontuário, sendo recodificado como perdido). Dos 536 pacientes inicialmente classificados como perdidos de seguimento, após a revisão dos prontuários, 109 foram recodificados como vivos, 160 haviam morrido (recodificados como óbito) e 267 ainda não tinham comparecido ao serviço nos 5 anos de seguimento, mantendo o status de perdido de seguimento. Portanto, o status dos pacientes no banco de dados final ficou caracterizado por 770 vivos, 992 óbitos e 522 perdidos de seguimento (Tabela 1).

Tabela 1 - Número e porcentagem de pacientes segundo status de seguimento inicial e final, RHC, 1994

Status original		Status final			Total	
		Vivo	Óbito	Perdido de seguimento		
				Nº	Nº	Nº
Vivo		661	201	255	1117	48,9
Óbito		-	631	-	631	27,6
Perdido de seguimento		109	160	267	536	23,5
Total	Nº	770	992	522	2284	100,0
	%	33,7	43,4	22,9	100,0	

Na Tabela 1, nota-se que no banco de dados original, 27,6% dos pacientes haviam falecido, enquanto que o banco de dados final (após a revisão/atualização da data do último comparecimento ao serviço), a proporção de óbitos foi de 43,4% uma diferença de 76% a mais de óbitos em 5 anos de seguimento.

4.2 Simulação das bases de dados

A base de dados padrão, já descrita, continha 770 pacientes vivos, 992 óbitos e 522 perdidos de seguimento (Tabela 2).

A divisão desta base de dados, conforme o status final do paciente resultou nas seguintes bases de dados:

1. com os pacientes considerados vivos até 5 anos (770 pacientes);
2. com os pacientes que foram a óbito em 5 anos (992 pacientes);
3. com os pacientes perdidos de seguimento ao longo de 5 anos (522 pacientes).

4.2.1 Perdas informativas nos óbitos

A segunda parte da Tabela 2 apresenta a formação das 10 bases de dados com perdas informativas nos óbitos. Nestas bases a quantidade de vivos era de 770 pacientes, a quantidade de perdas informativas variou de 50 a 502 e a de óbitos variou de 942 a 490 pacientes.

4.2.2 Perdas informativas nos vivos

A terceira parte da Tabela 2 apresenta a formação do banco de dados com perdas informativas nos vivos. Aos 770 pacientes vivos somou-se 522

perdidos de seguimento, resultando em 1292 pacientes vivos ao final do seguimento (censura) e 992 óbitos.

4.2.3 Perdas não informativas

A quarta parte da Tabela 2 apresenta o grupo das 10 bases de dados com perdas não informativas. Estas bases contêm menos 5 a 50% de pacientes comparadas com a base de dados padrão, pois todas foram geradas a partir de sorteio de 5 a 50% do total de pacientes.

4.2.4 Sem perdas anuais de 1994 a 1999

Por fim, na última parte da Tabela 2, estão apresentadas as formações dos 6 bancos de dados sem as perdas anuais (1994 a 1999). Vê-se que no primeiro ano foram excluídos 163 pacientes que deixaram de se apresentar ao serviço em 1994, 158 pacientes não se apresentaram, em 1995. Nos anos de 1996, 1997, 1998, e 1999 não compareceram ao serviço, respectivamente, 63, 49, 48 e 41 pacientes.

Tabela 2 – Nome, descrição geral e número de observações dos 28 bancos de dados gerados

Tipo de Banco	Nome do Banco	Descrição	Vivos	Óbitos	Perdidos	Amostra Final
Iniciais	Total	original com total de pacientes	770	992	522	2284
Perdas informativas nos óbitos	vvtt_ob5	Total de vivos com 5% de perda nos óbitos	770	942	50	1762
	vvtt_ob10	Total de vivos com 10% de perda nos óbitos	770	899	93	1762
	vvtt_ob15	Total de vivos com 15% de perda nos óbitos	770	817	175	1762
	vvtt_ob20	Total de vivos com 20% de perda nos óbitos	770	792	200	1762
	vvtt_ob25	Total de vivos com 25% de perda nos óbitos	770	474	245	1762
	vvtt_ob30	Total de vivos com 30% de perda nos óbitos	770	689	303	1762
	vvtt_ob35	Total de vivos com 35% de perda nos óbitos	770	635	357	1762
	vvtt_ob40	Total de vivos com 40% de perda nos óbitos	770	596	396	1762
	vvtt_ob45	Total de vivos com 45% de perda nos óbitos	770	547	445	1762
	vvtt_ob50	Total de vivos com 50% de perda nos óbitos	770	490	502	1762
Perdas informativas em vivos	Vvob_sem perda	perdidos foram recodificados como vivos	1292	992	-	2284
Perdas não informativas	perda_5	Total com 5% de perda aleatória	732	932	501	2165
	perda_10	Total com 10% de perda aleatória	693	880	472	2045
	perda_15	Total com 15% de perda aleatória	663	849	439	1951
	perda_20	Total com 20% de perda aleatória	622	787	426	1835
	perda_25	Total com 25% de perda aleatória	592	739	405	1736
	perda_30	Total com 30% de perda aleatória	532	723	371	1626
	perda_35	Total com 35% de perda aleatória	481	626	341	1448
	perda_40	Total com 40% de perda aleatória	475	593	317	1385
	perda_45	Total com 45% de perda aleatória	420	540	285	1245
	perda_50	Total com 50% de perda aleatória	375	464	249	1088
Sem perdas ao longo do seguimento	perda_1	Sem perda em 1994	770	992	359	2121
	perda_2	Sem perda até 1995	770	992	201	1963
	perda_3	Sem perda até 1996	770	992	138	1900
	perda_4	Sem perda até 1997	770	992	89	1851
	perda_5	Sem perda até 1998	770	992	41	1803
	perda_6	Sem perda até 1999	770	992	-	1762

4.3 Análise das perdas

Dos 2284 pacientes, 77% continuaram vivos ou foram a óbito e 23% foram perdidos de seguimento no período de 5 anos de seguimento. A maior porcentagem de perda ocorreu no primeiro ano de seguimento (1994) chegando a 163 pacientes, em 1995 deixaram de se apresentar ao serviço 158 pacientes, em 1996 não se apresentaram 63 pacientes, 49 pacientes em 1997, 48 pacientes em 1998 e 41 em 1999. Portanto os percentuais de perda foram maiores nos dois primeiros anos de seguimento.

A Tabela 3 mostra a distribuição dos pacientes segundo a situação de perdido de seguimento ou não (vivo ou óbito) para cada variável analisada, nesta tabela não foram considerados os pacientes com dados ignorados.

Não houve diferença estatisticamente significativa nas proporções dos perdidos com relação às variáveis sexo ($p=0,57$) e cor ($p=0,69$). Houve maior proporção de perda entre aqueles pacientes com ensino médio ou superior em relação aos de ensino fundamental (26,17% X 20,97%; $p=0,01$). A proporção de perdidos de seguimento foi maior entre os pacientes admitidos por convênio ou particulares (27,85% X 18,85%; $p<0,05$). Os pacientes com estágio clínico um ou dois também tiveram maior proporção de perda em relação aos pacientes com estágio clínico de três ou quatro (25,83% X 18,92%; $p<0,05$). Houve maior proporção de perda em pacientes com idade entre 15 e 39 anos e entre os mais idosos (25,93% e 25,15%; $p=0,03$).

Tabela 3 – Número e porcentagem de pacientes segundo status final e características demográficas e clínicas, RHC, 1994.

Variável	Status Final						p *
	Vivo ou óbito		Perdido		Total		
	N	%	N	%	N	%	
<u>Sexo</u>							
Feminino	947	76,68	288	23,32	1235	100,00	0,57
Masculino	815	77,69	234	22,31	1049	100,00	
<u>Faixa etária</u>							
0-14	153	83,15	31	16,85	184	100,00	0,03**
15-39	240	74,07	84	25,93	324	100,00	
40-64	854	78,49	134	21,51	1088	100,00	
65 e mais	515	74,85	173	25,15	688	100,00	
<u>Cor ou raça</u>							
Branca ou amarela	1501	77,37	439	22,63	1940	100,00	0,69
Preta ou parda	259	76,40	80	23,60	339	100,00	
<u>Escolaridade</u>							
Médio ou superior	440	73,83	156	26,17	596	100,00	0,01
Até fundamental	1221	79,03	324	20,97	1545	100,00	
<u>Categoria de admissão</u>							
Convênio ou particular	733	72,15	283	27,85	1016	100,00	<0,05
Sus	1029	81,15	239	18,85	1268	100,00	
<u>Estádio Clínico</u>							
Até II	600	74,17	209	25,83	809	100,00	<0,05
III ou mais	943	81,08	220	18,92	1163	100,00	

(*) Teste χ^2 com 5% de significância e 1 grau de liberdade.

(**) Teste χ^2 com 5% de significância e 3 graus de liberdade.

Não foram considerados pacientes com dados ignorados.

Dos 536 pacientes perdidos de seguimento no banco de dados inicial, 267 (49,81%) mantiveram o status, 160 compareceram ao serviço e 109 foram a óbito. E esta falta de recuperação foi maior no primeiro ano de seguimento, isto é, dos 267 perdidos de seguimento no banco de dados final, 120 compareceram ao serviço para o tratamento em 1994 e não fizeram seguimento após esta data (Tabela 4).

Tabela 4 Número e porcentagem dos pacientes com status inicial de perdidos de seguimento segundo status final e ano de seguimento, RHC, 1994

Ano	Status Final				Total	
	Recuperados		Perdidos		N	%
	N	%	N	%		
1994	38	24,05	120	75,95	158	100,00
1995	57	43,18	75	56,82	132	100,00
1996	26	48,15	28	51,85	54	100,00
1997	18	46,15	21	53,85	39	100,00
1998	13	56,52	10	43,48	23	100,00
1999	117	90,00	13	10,00	130	100,00
Total	269	50,19	267	49,81	536	100,00

A Tabela 5 apresenta a distribuição do total dos pacientes segundo a recuperação do status na revisão de prontuários (levando-se em conta que 267 permaneceram perdidos de seguimento) e às variáveis de estudo. Não houve diferença estatisticamente significativa nas proporções de pacientes que foram recuperados segundo idade ($p=0,32$), cor/raça ($p=0,18$), escolaridade ($p=0,12$) e estágio clínico ($p=0,12$). A proporção de perdidos foi maior entre os pacientes do sexo feminino ($p=0,04$) e entre os admitidos no hospital por convênio ou particular ($p<0,01$).

Tabela 5 -Número e percentual de pacientes segundo a recuperação na análise do prontuário e características demográficas e clínicas , RHC, 1994.

Variável	Status Final						p *
	Recuperados		Perdidos		Total		
	N	%	N	%	N	%	
<u>Sexo</u>							
Feminino	1075	87,04	160	12,96	1235	100,00	0,04
Masculino	942	89,80	107	10,20	1049	100,00	
<u>Faixa etária</u>							
0-14	166	90,22	18	9,78	184	100,00	0,50**
15-39	279	86,11	45	13,89	324	100,00	
40-64	965	88,69	123	11,31	1088	100,00	
65 e mais	607	88,23	81	11,77	688	100,00	
<u>Cor /raça</u>							
Branca ou amarela	1720	88,66	220	11,34	1940	100,00	0,18
Preta ou parda	292	86,14	47	13,86	339	100,00	
<u>Escolaridade</u>							
Médio ou superior	519	87,08	77	12,92	596	100,00	0,12
Até fundamental	1382	89,45	163	10,55	1545	100,00	
<u>Categoria de admissão</u>							
Convênio ou particular	875	86,12	141	13,88	1016	100,00	<0,01
Sus	1142	90,06	126	9,94	1268	100,00	
<u>Estádio Clínico</u>							
Até II	706	87,27	103	12,73	809	100,00	0,12
III ou mais	1041	89,51	122	10,49	1163	100,00	

(*) Teste χ^2 com 5% de significância e 1 grau de liberdade.

(**) Teste χ^2 com 5% de significância e 3 graus de liberdade.

Não foram considerados pacientes com dados ignorados.

4.4 Caracterização da amostra

Do total de pacientes, 1235 eram mulheres (54%) e 1049 homens (46%), a maioria era da cor branca ou amarela (85%). A idade média foi 52 anos (dp=20,3 anos), a idade mediana 57 anos e 30% possuíam 65 anos ou mais na data do diagnóstico. Do total, 596 (28%) estudaram além da 8ª série do primeiro grau. Mais da metade (59%) foram classificados com estadios clínicos III ou IV da doença (Tabela 6).

Tabela 6 – Número e distribuição percentual dos pacientes segundo características demográficas e clínicas RHC, 1994.

Variável	N	%
<u>Sexo</u>		
Feminino	1235	54,07
Masculino	1049	45,93
Total	2284	100,00
<u>Faixa etária</u>		
0-14	184	8,06
15-39	324	14,19
40-64	1088	47,64
65 e mais	688	30,12
Total	2284	100,00
<u>Cor/raça*</u>		
Branca ou amarela	1940	85,13
Preta ou parda	339	14,87
Total	2279	100,00
<u>Escolaridade*</u>		
Médio ou superior	596	27,84
Até fundamental	1545	72,16
Total	2141	100,00
<u>Categoria de admissão</u>		
Convênio ou particular	1016	44,48
Sus	1268	55,52
Total	2284	100,00
<u>Estadio Clínico*</u>		
Até II	809	41,02
III ou IV	1163	58,98
Total	1972	100,00

(*) Não foram considerados pacientes com dados ignorados.

4.5 Análise da probabilidade de sobrevida acumulada e HR no total dos pacientes

A Tabela 7 apresenta os valores das probabilidades de sobrevida acumuladas até o primeiro, terceiro e quinto ano de seguimento. A probabilidade de sobrevida acumulada para o total de pacientes foi de 75% para o primeiro ano de seguimento, 58% no terceiro e 50% no quinto ano de

seguimento. O sexo feminino teve maior probabilidade de sobrevida ao longo do seguimento ($p<0,05$). A idade, a escolaridade e o estágio clínico também tiveram diferenças significativas nas probabilidades de sobrevida, segundo suas categorias. Nota-se que os valores das probabilidades de sobrevida acumulada da categoria de estágio clínico I ou II apresentaram valores mais de duas vezes superior aos da categoria III ou IV após o terceiro ano de seguimento. Não houve diferença estatisticamente significativa nas probabilidades de sobrevida acumuladas segundo a cor/raça do paciente.

Tabela 7 – Probabilidade de sobrevida acumulada do total de pacientes segundo as características demográficas e clínicas até o primeiro, terceiro e quinto ano de seguimento, RHC, 1994.

Variável		Tempo			Teste <i>log rank</i> (p)
		1º ano	3º ano	5º ano	
Sexo	Feminino	0,82	0,66	0,58	<0,05
	Masculino	0,67	0,48	0,40	
Faixa Etária	0-14	0,80	0,67	0,64	<0,05
	15-39	0,84	0,71	0,64	
	40-64	0,77	0,59	0,51	
	65 e mais	0,66	0,46	0,38	
Cor/Raça	Branca ou amarela	0,75	0,58	0,51	0,32
	Preta ou Parda	0,74	0,54	0,47	
Categoria	Convênio ou Particular	0,78	0,63	0,56	<0,05
	Sus	0,73	0,55	0,47	
Escolaridade	Médio ou superior	0,80	0,65	0,59	<0,05
	Até fundamental	0,73	0,55	0,47	
Estádio Clínico	I+ II	0,94	0,85	0,79	<0,05
	III + IV	0,62	0,38	0,30	
Total	-	0,75	0,58	0,50	-

A força de mortalidade foi maior para os pacientes com estadiamento III ou IV ($HR=5,41$; $p<0,05$) quando comparado aos pacientes com menor estágio clínico. A força de mortalidade foi maior entre os homens

comparados às mulheres (HR=1,73; $p<0,05$). Os pacientes com mais de 64 anos apresentaram maior força de mortalidade (HR=2,15; $p<0,05$) comparados com os pacientes mais jovens. Os pacientes com menor escolaridade apresentaram maior força de mortalidade que os de maior escolaridade (HR=1,42; $p<0,05$). A categoria de admissão também apresentou maior força de mortalidade para os pacientes admitidos na categoria SUS (HR=1,30; $p<0,05$). A cor/raça do paciente não apresentou diferença na força de mortalidade entre as categorias analisadas ($p>0,05$) (Tabela 8).

Tabela 8 – Razão das funções de riscos segundo as características demográficas e clínicas, RHC, 1994.

Variável		HR	IC _{95%} (HR)	
Sexo	Feminino	1		
	Masculino	1,73	1,52	1,96
Faixa Etária	0-14	1		
	15-39	0,95	0,69	1,31
	40-64	1,43	1,09	1,86
	65 e mais	2,15	1,64	2,82
Cor/Raça	Branca ou amarela	1		
	Preta ou Parda	1,10	0,92	1,30
Categoria de Admissão	Convênio ou Particular	1		
	Sus	1,30	1,14	1,47
Escolaridade	Médio ou superior	1	-	-
	Até fundamental	1,42	1,22	1,66
Estádio Clínico	I + II	1	-	-
	III + IV	5,41	4,52	6,47

4.6 Análise do efeito das perdas nas probabilidades de sobrevivência acumuladas nas bases de dados

Nos Anexos de 4 a 17 estão apresentadas as tabelas com o detalhamento das análises feitas segundo Kaplan-Meier.

4.6.1 Total de pacientes

Houve um aumento significativo na probabilidade de sobrevida acumulada quando o tempo de seguimento foi até um ano (0,0028 para cada porcentagem de perda informativa), três anos (0,0041 para cada porcentagem de perda informativa) ou cinco anos (0,0041 para cada porcentagem de perda informativa). Observa-se que, quando as perdas são informativas, os percentuais de aumento nas probabilidades de sobrevida no primeiro ano variaram de 3,77% (até 10% de perda) a 19,46% (até 50% de perda). Houve um aumento destes valores quando o tempo de seguimento passa para 3 ou 5 anos (Tabela 9).

Quando as perdas foram não informativas, não houve diferença significativa nas probabilidades de sobrevida acumuladas e as variações percentuais nas probabilidades de sobrevida foram de 0,27% (com 10% de perda no primeiro ano de seguimento) a 2,20% (com 50% de perda no quinto ano de seguimento).

A Figura 2 apresenta estes resultados. Quando as perdas foram informativas, no primeiro ano, a probabilidade de sobrevida acumulada inicial é de 0,70 (momento zero sem perdas) passando para 0,83 (com 50% de perdas geradas); no terceiro ano, iniciou em 0,51 passando para 0,71 e no quinto ano, de 0,44 para 0,67.

A Figura 3 mostra as curvas de Kaplan Meier da base de dados padrão e das bases de dados com 50% de perda informativa nos óbitos e não informativas.

Tabela 9 – Modelos de regressão linear simples segundo as perdas de 5 a 50% e variação percentual das probabilidades de sobrevida acumuladas no total de pacientes, RHC, 1994

Tipo de Perda	Ano	β_0	β_1	r^2	p	Variação percentual			
						10%*	15%**	20%***	50%****
Informativa	1	0,6979	0,0028	0,99	< 0,01	3,77	5,96	7,06	19,46
	3	0,5088	0,0041	0,99	< 0,01	6,16	12,19	15,08	38,97
	5	0,4416	0,0041	0,99	< 0,01	9,77	17,24	20,28	53,43
Não informativa	1	0,7520	0,0000	0,01	0,43	0,27	-0,25	-0,20	0,39
	3	0,5766	0,0000	0,00	0,36	0,89	-0,31	0,36	1,58
	5	0,5019	0,0000	0,00	0,49	0,78	0,16	0,72	2,20

(*) com 10% de perda (**) com 15% de perda (***) com 20% de perda (****) com 50% de perda

Figura 2 – Gráficos das probabilidades de sobrevida observadas segundo o total de pacientes até o primeiro, terceiro e quinto ano de seguimento nos bancos gerados com perdas nos óbitos e aleatórias, RHC, 1994.

a – 3 Estimativas das probabilidades de sobrevivência para 4 bases de dados, RHC, 1994

4.6.2 Variável sexo

Houve um aumento significativo na probabilidade de sobrevida acumulada, para cada unidade percentual de perda informativa, para ambos os sexos no primeiro (0,0021 e 0,0035 para o sexo feminino e masculino respectivamente), terceiro (0,0036 e 0,0044 para o sexo feminino e masculino respectivamente) ou quinto ano de seguimento (0,0041 e 0,0047 para o sexo feminino e masculino respectivamente).

Nos pacientes do sexo feminino, observa-se que, quando as perdas são informativas, o percentual de aumento nas probabilidades de sobrevida no primeiro ano variou de 1,04% (com 10% de perda) a 12,78% (com 50% de perda), enquanto que para o sexo masculino, estes valores variaram de 7,61% (com 10% de perda) a 28,75% (com 50% de perda).

Quando o tempo de seguimento passa para 3 ou 5 anos, estes valores apresentam um aumento, chegando a atingir uma variação de 74,29% (para os pacientes do sexo masculino com 50% de perda no 5º ano de seguimento) com relação aos valores iniciais sem perda (Tabela 10).

Quando as perdas foram não informativas, os valores das probabilidades de sobrevida acumuladas não apresentaram variação significativa, como pode ser observado na Tabela 10, onde os coeficientes de regressão (β_1) dos modelos de regressão linear ajustados apresentaram valores nulos de variação para cada percentual de perda não informativa. Seguindo estes resultados, observa-se que os valores das variações percentuais nas probabilidades de sobrevida acumuladas foram menores que na perda informativa, atingindo o valor máximo de 2,50% (no quinto ano de seguimento com 50% de perdas geradas) para ambos os sexos.

A Figura 4 apresenta estes resultados e nota-se que a probabilidade acumulada de sobrevida foi maior para os pacientes do sexo feminino em todos os períodos de seguimento.

A Figura 5 mostra as curvas de Kaplan-Meier das bases de dados padrão e das bases de dados com 50% de perda informativa nos óbitos e não informativa.

Tabela 10 - Modelos da regressão linear simples segundo as perdas de 5 a 50% e variação percentual das probabilidades de sobrevivência acumuladas até o primeiro, terceiro e quinto ano de seguimento segundo sexo, RHC,1994

Tipo de Perda	Ano	Categoria	Estimadores da regressão				Variação percentual			
			b ₀	b ₁	r ²	p	10%*	15%**	20%***	50%***
informativa	1	Feminino	0,7797	0,0021	0,96	< 0,01	1,04	5,10	4,92	12,78
		Masculino	0,6026	0,0035	0,98	< 0,01	7,61	7,05	9,99	28,75
	3	Feminino	0,6010	0,0036	0,98	< 0,01	2,91	10,13	11,75	29,31
		Masculino	0,4017	0,0044	0,99	< 0,01	10,86	15,20	19,99	53,26
	5	Feminino	0,5320	0,0041	0,99	< 0,01	5,33	13,94	15,64	40,46
		Masculino	0,3364	0,0047	0,99	< 0,01	16,72	22,57	27,63	74,29
Não informativa	1	Feminino	0,8211	-0,0001	0,00	0,53	0,40	0,64	0,15	-0,12
		Masculino	0,6705	0,0002	0,11	0,31	0,07	-0,28	-0,80	-0,68
	3	Feminino	0,6579	0,0000	0,01	0,81	0,24	0,84	-0,17	1,42
		Masculino	0,4801	0,0001	0,07	0,81	1,31	1,02	-0,48	-2,01
	5	Feminino	0,5848	0,0000	0,00	0,86	1,59	1,10	-0,34	2,09
		Masculino	0,4028	0,0000	0,00	0,99	0,88	0,25	1,14	-2,51

(*) com 10% de perda (**) com 15% de perda (***) com 20% de perda (****) com 50% de perda

Figura 4 – Gráficos das probabilidades de sobrevida acumulada segundo o sexo até o primeiro, terceiro e quinto ano de seguimento nos bancos gerados com perdas nos óbitos e aleatórias , Registro de Câncer do Hospital AC Camargo , 1994.

Perda informativa

Perda não informativa

Figura – 5 Estimativas das probabilidades de sobrevivência para 4 bases de dados.

4.6.3 Variável idade

As probabilidades de sobrevida acumuladas apresentaram aumento significativo para cada unidade percentual de perda informativa, considerando o primeiro ano (0,0030, 0,0017, 0,0025 e 0,0024) para os pacientes com idades até 14, de 15 a 39, de 40 a 64 anos e 65 anos ou mais respectivamente, terceiro ano (de 0,0041, 0,0032, 0,0036 e 0,0050) para os pacientes com as idades até 14, de 15 a 39, de 40 a 64 anos e 65 anos ou mais anos respectivamente e quinto (de 0,0043, 0,0037, 0,0042 e 0,0036) para os pacientes com as idades até 14, de 15 a 39, de 40 a 64 anos e 65 anos ou mais respectivamente.

Observa-se que o percentual de aumento nas probabilidades de sobrevida considerando a idade dos pacientes, no primeiro ano de seguimento, variou de 1,51 (até 10% de perda) a 17,97 (até 50% de perda) para os pacientes com menos de 15 anos; de 1,50 (até 10% de perda) a 32,01 (até 50% de perda) para os pacientes com idades entre 15 e 49 anos e de 3,72 (até 10% de perda) a 29,22 (até 50% de perda) para os pacientes mais velhos. Estes valores foram maiores quando o tempo de seguimento passa para três ou cinco anos em todas faixas etárias consideradas (Tabela 11).

Quando as perdas foram não informativas, os valores das probabilidades de sobrevida acumuladas não apresentaram variação estatisticamente significativa. Observa-se que, com este tipo de perdas a maior variação percentual nas probabilidades de sobrevida acumuladas, foi 5,02% (para os pacientes mais de 64 anos de idade com 50% de perda no primeiro ano de seguimento) e a menor 0,25 (para os pacientes de 15 a 39 anos de idade com 10% de perda no primeiro ano de seguimento) (Tabela 11).

A Figura 6 apresenta estes resultados, verifica-se que os pacientes mais jovens apresentaram maior probabilidade de sobrevida que os mais velhos.

A Figura 7 mostra as curvas de Kaplan Meier das bases de dados padrão e das bases de dados com 50% de perda informativa nos óbitos e não informativas.

Tabela 11 – Modelos de regressão linear simples segundo as perdas de 5 a 50% e variação percentual das probabilidades de sobrevivência acumuladas até o primeiro, terceiro e quinto ano de seguimento segundo idade, RHC, 1994

Tipo de Perda	Ano	Categoria	β_0	β_1	r^2	p	Variação percentual			
							10%*	15%**	20%***	50%****
Informativa	1	0 - 14	0,7560	0,0030	0,96	< 0,01	1,51	3,65	4,05	17,97
		15 - 39	0,7973	0,0017	0,94	< 0,01	2,80	1,86	3,82	10,78
		40 - 64	0,7303	0,0025	0,99	< 0,01	3,57	5,74	6,66	17,40
		65 e mais	0,5934	0,0020	0,97	< 0,01	5,45	9,50	10,74	29,22
	3	0 - 14	0,6099	0,0041	0,94	< 0,01	1,50	8,57	8,76	32,01
		15 - 39	0,6392	0,0032	0,98	< 0,01	5,26	3,69	6,97	22,75
		40 - 64	0,5380	0,0036	0,99	< 0,01	6,49	13,01	16,47	35,78
		65 e mais	0,3681	0,0050	0,98	< 0,01	7,74	18,11	19,38	60,33
	5	0 - 14	0,5829	0,0043	0,92	< 0,01	3,72	8,90	10,27	37,82
		15 - 39	0,5757	0,0037	0,97	< 0,01	6,75	5,11	8,15	32,26
		40 - 64	0,4609	0,0042	0,98	< 0,01	11,53	17,84	22,61	50,21
		65 e mais	0,2952	0,0053	0,98	< 0,01	10,52	28,71	28,01	85,43
Não informativa	1	0 - 14	0,7940	-0,0002	0,04	0,53	-5,12	0,06	-6,42	-4,08
		15 - 39	0,8438	-0,0003	0,46	0,02	-0,25	0,93	-1,25	-1,96
		40 - 64	0,7596	0,0004	0,20	0,17	0,36	0,30	0,39	-0,09
		65 e mais	0,6413	0,0007	0,21	0,17	0,86	-1,31	0,45	-0,67
	3	0 - 14	0,6679	-0,0004	0,08	0,40	-0,91	1,49	-1,88	0,40
		15 - 39	0,7155	-0,0006	0,53	0,01	-0,37	1,35	-0,59	-4,04
		40 - 64	0,5758	0,0003	0,33	0,11	1,05	-0,15	1,17	0,02
		65 e mais	0,4364	0,0008	0,11	0,26	1,67	-2,04	0,13	2,70
	5	0 - 14	0,6385	0,0004	0,06	0,48	-0,92	1,77	-2,30	-0,77
		15 - 39	0,6505	-0,0003	0,16	0,21	-0,36	2,24	0,47	-1,63
		40 - 64	0,4999	0,0001	0,78	0,01	0,75	-1,00	1,04	-2,21
		65 e mais	0,3568	0,0009	0,06	0,34	0,13	-1,39	-0,34	5,02

(*) com 10% de perda (**) com 15% de perda (***) com 20% de perda (****) com 50% de perda

Figura 6 – Gráficos das probabilidades de sobrevida acumulada segundo a idade até o primeiro, terceiro e quinto ano de seguimento nos bancos gerados com perdas nos óbitos e aleatórias, RHC, 1994.

Figura – 7 Estimativas das probabilidades de sobrevivência para 4 bases de dados, RHC,1994

4.6.4 Variável Cor/raça

Observa-se que, quando as perdas foram informativas, houve um aumento significativo nas probabilidades acumuladas de sobrevida, para cada unidade percentual de perda informativa, para os pacientes de todas as categorias da variável cor/raça quando o tempo considerado foi um ano (0,0027 nos pacientes de cor/raça branca ou amarela e 0,0033 para os pacientes de cor/raça preta ou parda), 3 anos (0,0040 nos pacientes de cor/raça branca ou amarela e 0,0044 para os pacientes de cor/raça preta ou parda) ou 5 anos (0,0044 nos pacientes de cor/raça branca ou amarela e 0,0047 para os pacientes de cor/raça preta ou parda).

Observa-se que o percentual de aumento nas probabilidades de sobrevida variou de 7,16% (até 10% de perda) a 60,76 (até 50% de perda) com cinco anos de seguimento, para os pacientes de cor/raça preta ou parda (Tabela 12).

Quando as perdas são não informativas não houve diferença significativa nas probabilidades de sobrevida acumulada dos pacientes com relação a variável cor/raça. Deste modo a maior variação percentual apresentada na probabilidade de sobrevida acumulada foi -12,07% (para a cor/raça branca ou amarela no quinto ano de seguimento com até 20% de perda).

A Figura 8 apresenta estes resultados. Nota-se que as probabilidades de sobrevida acumuladas dos pacientes de cor/raça branca ou amarela apresentam valores próximos aos pacientes de cor/raça preta ou parda ao longo do período de seguimento.

A Figura 9 mostra as curvas de Kaplan Meier das bases de dados padrão e das bases de dados com 50% de perda informativa nos óbitos e não informativas.

Tabela 12 Modelos de regressão linear simples segundo as perdas de 5 a 50% e variação percentual das probabilidades de sobrevivência acumuladas até o primeiro, terceiro e quinto ano de seguimento segundo cor/raça, RHC, 1994

Tipo de Perda	Ano	Categoria	β_0	β_1	r^2	p	Variação percentual			
							10%*	15%**	20%***	50%***
informativa	1	Branca ou Amarela	0,6999	0,0027	0,98	< 0,01	3,78	5,87	7,04	18,63
		Preta ou Parda	0,6839	0,0033	0,87	< 0,01	3,81	6,58	7,34	24,53
	3	Branca ou Amarela	0,5160	0,0040	0,98	< 0,01	6,32	12,30	15,06	37,93
		Preta ou Parda	0,4691	0,0044	0,95	< 0,01	5,28	10,88	15,43	44,53
	5	Branca ou Amarela	0,4488	0,0044	0,99	< 0,01	10,22	17,50	20,41	52,14
		Preta ou Parda	0,4000	0,0047	0,96	< 0,01	7,16	14,88	19,62	60,76
Não informativa	1	Branca ou Amarela	0,7510	0,0002	0,09	0,37	-8,05	-0,31	-8,50	-7,66
		Preta ou Parda	0,7550	-0,0006	0,29	0,09	0,51	-0,08	0,18	-0,91
	3	Branca ou Amarela	0,5812	0,0007	0,03	0,60	-11,08	0,09	-11,67	-10,83
		Preta ou Parda	0,5445	-0,0004	0,04	0,55	1,04	-2,99	1,50	5,48
	5	Branca ou Amarela	0,5100	0,0001	0,01	0,73	-11,86	0,41	-12,07	-11,17
		Preta ou Parda	0,4657	-0,0001	0,00	0,87	-1,85	-1,97	-0,62	5,06

(*) com 10% de perda (**) com 15% de perda (***) com 20% de perda (****) com 50% de perda

Figura 8 – Gráficos das probabilidades de sobrevida acumulada segundo o cor/raça até o primeiro, terceiro e quinto ano de seguimento nos bancos gerados com perdas nos óbitos e aleatórias , RHC , 1994.

Perda informativa

Perdas não informativa

— 1º ano — 3º ano — 5º ano

Figura – 9 Estimativas das probabilidades de sobrevivência para 4 bases de dados, RHC,1994

4.6.5. Variável Escolaridade

Houve um aumento significativo nas probabilidades de sobrevida acumulada dos pacientes de todos os níveis de escolaridade acumulada, para cada unidade percentual de perda informativa, considerando o primeiro ano (0,0025 para os pacientes com ensino médio ou superior e 0,0029 para os pacientes que estudaram até o fundamental), o terceiro ano (0,0035 para os pacientes com ensino médio ou superior e 0,0043 para os pacientes que estudaram até o fundamental) ou o quinto ano de seguimento (0,0040 para os pacientes com ensino médio ou superior e 0,0047 para os pacientes que estudaram até o fundamental).

Observa-se que, no primeiro ano o percentual de aumento nas probabilidades de sobrevida variou de 3,42% (até 10% de perda) a 17,22% (até 50% de perda) para os pacientes com ensino médio ou superior e de 3,82 (até 10% de perda) a 20,55 (até 50% de perda) para os pacientes que estudaram até o ensino fundamental (Tabela 13).

Quando as perdas foram não informativas, não houve variação significativa nas probabilidades de sobrevida para os pacientes de qualquer nível de escolaridade. A variação percentual nas probabilidades de sobrevida acumuladas apresentou valor máximo de 4% (para os pacientes com ensino médio ou superior com até 50% de perda em 5 anos de seguimento).

A Figura 10 apresenta estes resultados e nota-se que os pacientes que estudaram até o ensino médio ou superior apresentaram probabilidade de sobrevida acumulada maior que os pacientes de menor escolaridade.

A Figura 11 mostra as curvas de Kaplan-Meier da base de dados padrão e das bases de dados com 50% de perda informativa nos óbitos e não informativas.

Tabela 13 – Modelos de regressão linear simples segundo as perdas de 5 a 50% e variação percentual das probabilidades de sobrevida acumuladas até o primeiro, terceiro e quinto ano de seguimento segundo escolaridade, RHC, 1994

Tipo de Perda	Ano	Categoria	β_0	β_1	r^2	p	Variação percentual			
							10%*	15%**	20%***	50%****
Informativas	1	médio ou superior	0,7396	0,0025	0,99	< 0,01	3,42	4,53	6,41	17,22
		até fundamental	0,6837	0,0029	0,99	< 0,01	3,82	6,58	7,34	20,55
	3	médio ou superior	0,5823	0,0035	0,98	< 0,01	5,28	7,65	12,96	30,13
		até fundamental	0,4836	0,0043	0,98	< 0,01	6,50	13,85	15,81	43,34
	5	médio ou superior	0,5241	0,0040	0,99	< 0,01	8,04	10,45	16,39	43,15
		até fundamental	0,4120	0,0047	0,98	< 0,01	10,35	19,88	21,75	58,88
Não informativa	1	médio ou superior	0,7932	0,0002	0,36	0,05	1,82	-0,08	1,76	3,46
		até fundamental	0,7357	-0,0005	0,06	0,83	0,71	-0,35	0,26	-0,07
	3	médio ou superior	0,6509	-0,0001	0,01	0,75	1,34	-0,11	0,97	2,37
		até fundamental	0,5472	0,0001	0,00	0,96	-0,27	0,92	0,05	0,62
	5	médio ou superior	0,5852	0,0002	0,14	0,25	-0,14	0,58	0,19	4,00
		até fundamental	0,4692	0,0000	0,00	0,92	-0,04	-0,42	0,49	0,17

(*) com 10% de perda (**) com 15% de perda (***) com 20% de perda (****) com 50% de perda

Figura 10 – Gráficos das probabilidades de sobrevivência acumulada segundo a escolaridade até o primeiro, terceiro e quinto ano de seguimento nos bancos gerados com perdas nos óbitos e aleatórias, RHC , 1994.

Perda informativa

Perdas não informativa

Figura – 11 Estimativas das probabilidades de sobrevivência para 4 bases de dados, RHC,1994

4.6.6 Variável categoria de admissão

Observa-se que, quando as perdas foram informativas, houve um aumento significativo nas probabilidades de sobrevida acumulada para os pacientes de todas as categorias de admissão, para cada unidade percentual de perda, quando considerado um ano (0,0027 nos pacientes admitidos por convênio ou particular e 0,0028 nos admitidos pelo SUS), três anos (0,0039 nos pacientes admitidos por convênio ou particular e 0,0028 nos admitidos pelo SUS) ou cinco anos (0,0044 nos pacientes admitidos por convênio ou particular e 0,0045 nos admitidos pelo SUS).

As probabilidades de sobrevida, considerando o primeiro ano de seguimento, variaram de 3,87% (até 10% de perda) a 17,95% (até 50% de perda) para os pacientes admitidos por convênio ou particular e de 3,73% (até 10% de perda) a 20,57% (até 50% de perda) para os pacientes SUS (Tabela 14).

Quando as perdas são não informativas não houve diferença significativa nas probabilidades de sobrevida acumulada dos pacientes com relação a variável categoria de admissão. Assim, a maior variação percentual apresentada na probabilidade de sobrevida acumulada foi 3,48% (para os pacientes admitidos pelo SUS no quinto ano de seguimento com até 50% de perda).

A Figura 12 apresenta estes resultados. Nota-se que as probabilidades de sobrevida acumuladas dos pacientes admitidos através de convênio ou particulares são maiores que as dos pacientes admitidos pelo SUS.

A Figura 13 mostra as curvas de Kaplan-Meier da base de dados padrão e das bases de dados com 50% de perda informativa nos óbitos e não informativas.

Tabela 14 Modelos de regressão linear simples segundo as perdas de 5 a 50% e variação percentual das probabilidades de sobrevida acumuladas até o primeiro, terceiro e quinto ano de seguimento segundo categoria de admissão, RHC, 1994

Tipo de Perda	Ano	Categoria	β_0	β_1	r^2	p	Variação percentual			
							10%*	15%**	20%***	50%****
Informativa	1	Particular ou Convenio Sus	0,7156	0,0027	0,97	< 0,01	3,87	5,02	7,06	17,95
			0,6854	0,0028	0,98	< 0,01	3,73	6,64	7,06	20,57
	3	Particular ou Convenio Sus	0,5431	0,0039	0,99	< 0,01	5,57	11,56	14,48	35,01
			0,4845	0,0042	0,99	< 0,01	6,62	12,63	15,58	42,00
	5	Particular ou Convenio Sus	0,4807	0,0044	0,98	< 0,01	8,53	15,56	18,93	47,15
			0,4137	0,0045	0,99	< 0,01	10,73	18,53	21,35	58,45
Não Informativa	1	Particular ou Convenio Sus	0,7776	0,0002	0,26	0,11	0,24	-0,40	-0,99	1,10
			0,7323	0,0000	0,01	0,78	0,12	-0,16	0,30	-0,30
	3	Particular ou Convenio Sus	0,6276	-0,0001	0,07	0,42	1,05	0,24	-0,79	-0,27
			0,5381	0,0000	0,02	0,66	-0,13	-0,84	0,57	2,46
	5	Particular ou Convenio Sus	0,5583	0,0000	0,00	0,83	1,04	-0,48	-0,27	0,04
			0,4598	0,0000	0,00	0,95	-0,28	0,59	2,46	3,48

(*) com 10% de perda (**) com 15% de perda (***) com 20% de perda (****) com 50% de perda

Figura 12 – Gráficos das probabilidades de sobrevida acumulada segundo a categoria até o primeiro, terceiro e quinto ano de seguimento nos bancos gerados com perdas nos óbitos e aleatórias, RHC 1994.

Perda informativa

Perdas não informativa

— 1º ano — 3º ano — 5º ano

Figura – 13 Estimativas das probabilidades de sobrevivência para 4 bases de dados, RHC, 1994

4.6.7 Variável estadio clínico

Houve um aumento significativo nas probabilidades de sobrevida acumuladas, para todos os estádios clínicos, para cada unidade percentual de perda informativa no primeiro (0,0008 e 0,0038 para o estadio clínico I ou II e III ou IV respectivamente), terceiro (0,0018 e 0,0048 para os pacientes com estadio clínico I ou II e III ou IV respectivamente) ou quinto ano de seguimento (0,0023 e 0,0049 para os pacientes com estadio clínico I ou II e III ou IV respectivamente).

Nos pacientes com estágio clínico I ou II, os percentuais de aumento nas probabilidades de sobrevida, no primeiro ano variaram de 0,88% (com 10% de perda) a 3,91% (com 50% de perda), enquanto que para os pacientes com estadio clínico III ou IV estes valores variaram de 6,37% (com 10% de perda) a 33,84% (com 50% de perda). Quando o tempo de seguimento passa para 3 ou 5 anos estes valores apresentam um aumento chegando a atingir uma variação de 93,23% (para os pacientes com estadio clínico III ou IV com 50% de perda no 5º ano de seguimento) com relação aos valores iniciais (Tabela 15).

Quando as perdas foram não informativas os valores das probabilidades de sobrevida acumuladas não apresentaram variação significativa como pode ser observado na Tabela 15, onde os coeficientes de regressão (β_1) dos modelos de regressão linear ajustados apresentaram valores nulos de variação para cada percentual de perda não informativa.

Seguindo estes resultados observa-se que os valores das variações percentuais nas probabilidades de sobrevida acumulada foram menores que nas perdas não informativas, atingindo o valor máximo de 2,20% (no quinto ano de seguimento com 50% de perdas geradas).

A Figura 14 apresenta estes resultados e observa-se que a probabilidade acumulada de sobrevida foi maior para os pacientes menos graves em todos os períodos de seguimento.

A Figura 15 mostra as curvas de Kaplan-Meier da base de dados padrão e das bases de dados com 50% de perda informativa nos óbitos e não informativa.

Tabela 15 – Modelos de regressão linear simples segundo as perdas de 5 a 50% e variação percentual das probabilidades de sobrevivência acumuladas até o primeiro, terceiro e quinto ano de seguimento segundo estadió clínico, RHC, 1994

Tipo de Perda	Ano	Categoria	β_0	β_1	r^2	p	Variação percentual			
							10%*	15%**	20%***	50%****
Informativa	1	até II III ou IV	0,9215	0,0008	0,90	< 0,01	0,88	0,88	1,75	3,91
			0,5559	0,0038	0,96	< 0,01	6,37	10,28	11,80	33,84
	3	até II III ou IV	0,8220	0,0018	0,95	< 0,01	1,28	3,45	4,96	10,03
			0,3122	0,0048	0,97	< 0,01	11,66	21,82	26,97	74,28
	5	até II III ou IV	0,7674	0,0023	0,95	< 0,01	3,68	6,58	7,68	17,00
			0,2462	0,0049	0,98	< 0,01	11,66	25,02	31,41	93,23
Não informativa	1	até II III ou IV	0,9386	-0,0001	0,61	0,03	-0,03	-0,04	-0,39	-0,68
			0,6161	0,0003	0,33	10,00	-0,85	-0,55	-0,90	2,13
	3	até II III ou IV	0,8527	0,0000	0,96	0,00	0,28	-	0,82	1,88
			0,3777	0,0001	0,40	0,08	-0,21	-0,77	-0,26	3,00
	5	até II III ou IV	0,7927	0,0000	0,85	0,00	-0,13	0,14	2,13	3,16
			0,3013	0,0002	0,12	0,25	1,20	0,79	-1,07	4,80

(*) com 10% de perda (**) com 15% de perda (***) com 20% de perda (****) com 50% de perda

Figura 14 – Gráficos das probabilidades de sobrevida acumulada segundo estadio até o primeiro, terceiro e quinto ano de seguimento nos bancos gerados com perdas nos óbitos e aleatórias , RHC, 1994.

Perda informativa

Perdas não informativa

— 1º ano — 3º ano — 5º ano

Figura – 15 Estimativas das probabilidades de sobrevida para 4 bases de dados, RHC,1994

4.7 Análise do efeito das perdas nas HR's

A Tabela 16 apresenta os resultados da análise das perdas nos valores das HR's em todos os bancos de dados. Quando as perdas foram informativas, houve acréscimo significativo nas HR das variáveis: sexo (0,0020 para cada percentual de perda nos óbitos comparando a sobrevivência dos pacientes do sexo feminino com as pacientes do sexo masculino) e idade (0,0031 para os pacientes com 15 a 39 anos e comparados aos mais jovens).

Quando as perdas foram não informativas não houve variação significativa nas HR em todas as variáveis consideradas na análise.

As variações percentuais apresentaram diversos comportamentos com aumentos (2,80% na HR da variável sexo com 10% de perda nos óbitos), quedas (0,56% de queda na mesma variável com 20% de perda) e posterior aumento (7,30% na mesma variável com 50% de perda).

As Figuras de 16 a 22 apresentam estes resultados e nota-se que o estágio clínico apresentou o maior valor de HR (valores próximos de 5) seguido pela variável sexo (valores próximos de 2).

Tabela 16 - Modelos de regressão linear simples segundo as perdas de 5 a 50% e variação percentual nas razões das funções de risco segundo as variáveis de estudo, RHC,1994

Variável	Tipo De Perda	β_0	β_1	r^2	p	Variação Percentual			
						10%*	15%**	20%***	50%****
Sexo	Informativa	1,7622	0,0019	0,26	0,11	-4,78	4,04	2,15	6,00
	Não informativa	1,7410	-0,0003	0,00	0,86	1,62	-0,20	4,71	-0,21
Faixa Etária	Informativa	0,9786	0,0038	0,35	0,06	-6,73	4,81	-4,81	8,00
		1,3882	0,0038	0,21	0,15	-2,80	-2,10	-8,39	10,77
		2,2463	0,0022	0,04	0,57	2,79	3,26	1,40	11,77
	Não informativa	0,9205	0,0002	0,00	0,93	-0,18	-0,93	-3,36	6,95
		1,4229	-0,0020	0,09	0,37	-3,55	4,54	6,77	5,51
		2,1297	-0,0033	0,11	0,42	-3,02	5,20	-5,05	-1,29
Cor/raça	Informativa	1,1200	-0,0011	0,24	0,12	1,82	2,73	0,91	-7,27
	Não informativa	1,0904	0,0010	0,05	0,52	2,04	1,90	-0,52	-4,69
Escolaridade	Informativa	1,3110	0,0002	0,00	0,88	2,56	-2,63	2,56	8,34
	Não informativa	1,3976	0,0010	0,06	0,45	-0,50	-1,74	-1,12	5,22
Categoria	Informativa	1,1725	0,0006	0,08	0,39	-1,14	0,04	2,48	0,04
	Não informativa	1,2950	0,0000	0,00	0,95	2,38	-0,57	-3,00	-2,75
Estádio	Informativa	4,0655	0,0017	0,04	0,38	0,50	3,48	-1,74	2,74
	Não informativa	5,4356	0,0000	0,00	0,99	-1,08	-0,24	9,12	8,31

(*) com 10% de perda (**) com 15% de perda (***) com 20% de perda (****) com 50% de perda

Figura 16 – Razão das funções de riscos segundo sexo, RHC , 1994.

Figura 17– Razão das funções de riscos segundo idade, RHC, 1994.

Figura 18 – Razão das funções de riscos segundo cor/raça,RHC , 1994.

Figura 19 – Razão das funções de riscos segundo escolaridade, RHC , 1994.

Figura 20 – Razão das funções de riscos segundo categoria, RHC , 1994.

Figura 21 – Razão das funções de riscos segundo estadio, RHC , 1994.

4.8 Análise do efeito das perdas dos anos de 1994 a 1999 na probabilidade acumulada de sobrevida no total dos pacientes

Nesta análise os pacientes perdidos de observação foram retirados do banco, ano a ano (de 1994 a 1999), não sendo considerados na análise. Nota-se que há uma queda nas probabilidades de sobrevida acumulada em todos os anos, para todas as variáveis estudadas. (Tabela 17).

Obviamente o efeito maior ocorre quando são considerados os percentuais de perdas ao longo dos anos. Vale lembrar os percentuais de perdas de seguimento foram de 7,14% no primeiro ano, 6,92% no segundo ano, 2,76% no terceiro, 2,10% no quarto e 1,80 no quinto ano de seguimento com relação aos 2284 pacientes do banco de dados inicial.

As Figuras 22 a 28 apresentam estes resultados e nota-se que a probabilidade de sobrevida acumulada para o sexo feminino passou de 0,82 (1994) para 0,78 (sem perdas até 1999). Nota-se que houve queda nas probabilidades de sobrevida acumulada e esta diminuição foi mais acentuada nos primeiros dois anos de seguimento.

Tabela 17 Variação percentual das probabilidades de sobrevida acumuladas das bases de dados sem perdas de 1994 a 1999 segundo as variáveis de estudo, RHC, 1994

TEMPO	CATEGORIA	$\Delta\%$ 0-1995	$\Delta\%$ 0-1996	$\Delta\%$ 0-1997	$\Delta\%$ 0-1998	$\Delta\%$ 0-1999
1	Total	-2,8	-3,9	-4,9	-6,0	-6,9
2		-3,0	-4,9	-6,9	-9,1	-11,0
3		-3,0	-4,9	-7,0	-9,4	-11,9
1	FEMININO	-2,0	-2,6	-3,3	-4,0	-4,7
	MASCULINO	-3,6	-5,9	-7,1	-8,7	-9,9
3	FEMININO	-2,2	-3,3	-4,8	-6,3	-7,8
	MASCULINO	-3,9	-7,4	-10,0	-13,3	-15,8
5	FEMININO	-2,2	-3,3	-4,8	-6,6	-8,5
	MASCULINO	-3,9	-7,4	-10,1	-13,8	-17,1
1	BRANCA OU AMARELA	-2,6	-3,8	-4,8	-5,9	-6,8
	PRETA OU PARDA	-3,4	-5,0	-5,4	-6,3	-7,2
3	BRANCA OU AMARELA	-2,8	-4,6	-6,8	-8,9	-10,8
	PRETA OU PARDA	-3,8	-6,4	-7,4	-9,4	-11,6
5	BRANCA OU AMARELA	-2,8	-4,6	-6,8	-9,4	-11,7
	PRETA OU PARDA	-3,8	-6,4	-7,4	-9,6	-12,3
1	PARTICULAR OU CONVÊNIO	-2,7	-4,0	-5,3	-6,7	-7,8
	SUS	-2,8	-3,9	-4,5	-5,2	-6,0
3	PARTICULAR OU CONVÊNIO	-2,8	-5,0	-7,6	-10,4	-12,5
	SUS	-3,1	-4,8	-6,1	-7,6	-9,3
5	PARTICULAR OU CONVÊNIO	-2,8	-5,0	-7,6	-10,8	-13,5
	SUS	-3,1	-4,8	-6,2	-7,9	-10,2
1	MAIS QUE FUNDAMENTAL	-2,2	-3,1	-4,4	-5,6	-6,6
	ATÉ FUNDAMENTAL	-3,0	-4,1	-4,9	-5,8	-6,6
3	MAIS QUE FUNDAMENTAL	-2,3	-3,6	-6,2	-8,4	-10,4
	ATÉ FUNDAMENTAL	-3,2	-5,2	-6,8	-8,8	-10,6
5	MAIS QUE FUNDAMENTAL	-2,3	-3,6	-6,2	-8,7	-11,1
	ATÉ FUNDAMENTAL	-3,2	-5,2	-6,9	-9,2	-11,5
1	I + II	-0,5	-0,8	-1,1	-1,5	-1,8
	III + IV	-4,8	-6,7	-7,7	-8,5	-9,6
3	I + II	-0,5	-1,0	-1,7	-2,7	-3,4
	III + IV	-5,4	-8,7	-11,3	-13,2	-16,0
5	I + II	-0,5	-1,0	-1,8	-3,0	-4,0
	III + IV	-5,4	-8,7	-11,4	-13,5	-17,3
1	DE 0 A 14 ANOS	-0,8	-1,4	-1,6	-2,5	-3,4
	DE 15 A 39 ANOS	-0,9	-1,8	-3,0	-4,1	-4,7
	DE 40 A 64 ANOS	-2,6	-2,6	-3,9	-5,2	-5,2
	DE 65 E MAIS ANOS	-4,6	-6,2	-7,7	-9,2	-10,8
3	DE 0 A 14 ANOS	-0,8	-1,8	-2,1	-3,7	-5,5
	DE 15 A 39 ANOS	-1,0	-2,4	-4,7	-7,0	-8,3
	DE 40 A 64 ANOS	-3,4	-5,1	-6,8	-13,6	-10,2
	DE 65 E MAIS ANOS	-4,4	-6,7	-11,1	-13,3	-15,6
5	DE 0 A 14 ANOS	-0,8	-1,8	-2,1	-3,9	-5,9
	DE 15 A 39 ANOS	-1,0	-2,4	-4,8	-7,5	-9,0
	DE 40 A 64 ANOS	-2,0	-4,0	-4,0	-6,0	-10,0
	DE 65 E MAIS ANOS	-2,7	-8,1	-10,8	-13,5	-18,9

Figura 22 – Gráficos das probabilidades de sobrevida acumulada segundo sexo até o primeiro, terceiro e quinto ano de seguimento nos bancos gerados sem perdas de 1994 a 1999, RHC, 1994.

Figura 23 – Gráficos das probabilidades de sobrevivência acumulada segundo idade até o primeiro, terceiro e quinto ano de seguimento nos bancos gerados sem perdas de 1994 a 1999, RHC, 1994.

Figura 24 – Gráficos das probabilidades de sobrevivência acumulada segundo raça até o primeiro, terceiro e quinto ano de seguimento nos bancos gerados sem perdas de 1994 a 1999. RHC . 1994.

Figura 25 – Gráficos das probabilidades de sobrevivência acumulada segundo escolaridade até o primeiro, terceiro e quinto ano de seguimento nos bancos gerados sem perdas de 1994 a 1999, RHC, 1994.

Figura 26 – Gráficos das probabilidades de sobrevida acumulada segundo categoria até o primeiro, terceiro e quinto ano de seguimento nos bancos gerados sem perdas de 1994 a 1999, RHC, 1994.

Figura 27 – Gráficos das probabilidades de sobrevivência acumulada segundo estadios até o primeiro, terceiro e quinto ano de seguimento nos bancos gerados sem perdas de 1994 a 1999, RHC, 1994.

Figura 28 – Gráficos das probabilidades de sobrevivência acumulada do total até o primeiro, terceiro e quinto ano de seguimento nos bancos gerados sem perdas de 1994 a 1999, RHC, 1994.

4.9 Análise do efeito das transformações das perdas de seguimento em vivo ao final do estudo na probabilidade acumulada de sobrevida e nas HR's no total dos pacientes

Nesta análise os 522 pacientes perdidos de observação no banco de dados original foram considerados vivos ao final de seguimento, simulando uma perda informativa nos vivos. Nota-se que há um aumento nas probabilidades de sobrevida acumulada no total dos pacientes, que passou de 0,75 para 0,77 no primeiro ano de seguimento gerando uma variação de 2,67%. Quando o tempo de seguimento aumenta, as diferenças se acentuam alcançando 14% com cinco anos de seguimento. Considerando as variáveis de estudo, o comportamento é semelhante, com aumentos nas probabilidades de sobrevida acumulada gerando diferenças que vão de 1,19% (idade de 15 a 39 anos com um ano de seguimento) a 30% (estádio clínico III ou IV com cinco anos de seguimento) (Tabela 18).

Tabela 18 – Probabilidade de sobrevida acumulada do total de pacientes segundo as características demográficas e clínicas até o primeiro, terceiro e quinto ano de seguimento com as perdas informativas nos vivos, RHC, 1994.

Variável		Ano			Teste <i>log rank</i> (p)	Variação Percentual*		
		1º	3º	5º		1º ano	3º ano	5º ano
Sexo	Feminino	0,83	0,70	0,64	<0,05	1,22	6,06	10,34
	Masculino	0,69	0,54	0,48		2,99	12,50	20,00
Faixa Etária	0-14 anos	0,81	0,69	0,67	<0,05	1,25	2,99	4,69
	15-39 anos	0,85	0,74	0,70		1,19	4,23	9,38
	40 –64 anos	0,79	0,63	0,57		2,60	6,78	11,76
	65 e mais anos	0,69	0,53	0,48		4,55	15,22	26,32
Cor/Raça	Branca ou amarela	0,77	0,63	0,60	0,32	2,67	8,62	17,65
	Preta ou Parda	0,76	0,60	0,55		2,70	11,11	17,02
Categoria de Admissão	Convenio ou Particular	0,80	0,67	0,63	<0,05	2,56	6,35	12,50
	Sus	0,74	0,58	0,52		1,37	5,45	10,64
Escolaridade	Médio ou superior	0,81	0,69	0,65	<0,05	1,25	6,15	10,17
	Até fundamental	0,75	0,60	0,54		2,74	9,09	14,89
Estádio Clínico	Até II	0,94	0,86	0,82	<0,05	0,00	14,12	3,80
	III + IV	0,64	0,45	0,39		3,23	18,42	30,00
Total	-	0,77	0,62	0,57	-	2,67	6,90	14,00

(*) comparado os valores do banco total padrão

A transformação dos perdidos de observação em vivos ao final do seguimento gerou diminuição nas HR's de 0,70% (na variável escolaridade) a 9,80% (variável estadiamento). A exceção da variável a categoria de admissão, que sofreu um aumento de 4,62% na HR (Tabela 19).

Tabela 19 – Razão das funções de riscos segundo as características demográficas e clínicas com perdas informativas nos vivos, RHC, 1994.

Variável		HR	IC _{95%} (HR)		Varição Percentual*
Sexo	Feminino	1	-	-	
	Masculino	1,67	1,47	1,90	-3,47
Faixa Etária	0-14	1	-		
	15-39	0,90	0,65	1,24	-5,26
	40 –64	1,33	1,02	1,74	-6,29
	65 e mais	1,83	1,40	2,40	-14,88
Cor/Raça	Branca ou amarela	1	-	-	
	Preta ou Parda	1,06	0,90	1,27	-3,64
Categoria de Admissão	Convenio ou Particular	1	-	-	
	Sus	1,36	1,20	1,55	4,62
Escolaridade	Médio ou superior	1	-	-	
	Até fundamental	1,41	1,21	1,65	-0,70
Estadio Clínico	até II	1	-	-	
	III + IV	4,88	4,09	5,84	-9,80

(*) comparado aos valores do banco total original

5. DISCUSSÃO

A origem da análise de dados de sobrevida é atribuída ao trabalho de John Graunt que, no início do século XVII, publicava em Londres, listas descrevendo os óbitos e suas causas. A partir dos dados de mortalidade, ele construiu uma tabela, que a partir de um grupo de 100 vidas, indicava o número de sobreviventes a cada intervalo de 10 anos entre as idades de 6 a 76 anos.

A primeira tábua de vida foi publicada por Edmond Halley em 1693, descrevendo as taxas de sobrevivência dos habitantes de uma cidade da Alemanha a pedido da Real Sociedade de Londres.

O aprimoramento das técnicas de análise de sobrevida se deu durante a Segunda Guerra Mundial, devido ao interesse em analisar a durabilidade e a confiabilidade do material bélico utilizado. Após este período, estes métodos começaram a ser empregados na área da saúde para analisar o tempo de vida de pacientes com câncer (LEE 1997).

Nas décadas de 50 e 60, apareceram as primeiras propostas de estimadores das probabilidades de sobrevida incorporando a censura. Uma delas foi o modelo não paramétrico proposto por Kaplan e Meier denominado estimador produto limite de Kaplan-Meier. Por este método, a divisão do tempo é sempre determinada por uma falha, por exemplo, o óbito. Através deste modelo, pode-se comparar o conjunto de curvas de sobrevida das diversas categorias de uma única variável independente.

Em 1972, D. R. Cox desenvolveu um modelo de regressão semi-paramétrico, também conhecido como modelo de riscos proporcionais de Cox. Com esta técnica, é possível avaliar o impacto de um conjunto de fatores prognósticos no tempo até a ocorrência do evento de interesse.

A disponibilidade destas duas técnicas de análise de sobrevida, na maior parte dos pacotes estatísticos e a facilidade de interpretação e apresentação de seus resultados fazem com que elas sejam o método estatístico de escolha dos estudos envolvendo análise de sobrevida

(HOSMER e LEMESHOW,1999). Apesar disto, muitos pesquisadores não levam em conta as suposições referentes aos tempos de seguimento incompletos, ou seja, às censuras à direita na análise dos dados.

Em epidemiologia, os ensaios clínicos e os estudos de coorte são os delineamentos que podem utilizar o tempo como variável de interesse. Nestes estudos, há o interesse em se verificar o efeito dos fatores de risco ou de fatores prognósticos no tempo de sobrevida de um indivíduo ou de um grupo, bem como definir as probabilidades ao longo do tempo.

Os ensaios clínicos, em geral, constituem-se de grupos menores que os estudos de coorte e pressupõem um acompanhamento mais detalhado dos pacientes. Uma das limitações apontadas por PEREIRA (1995) tanto para os estudos de coorte, quanto para os ensaios clínicos é a possibilidade de perdas de seguimento de alguns pacientes.

HENNEKENS e BURING (1987) afirmam que se a perda for grande em um estudo de coorte (entre 30 e 40%), os resultados do estudo podem estar seriamente prejudicados. E mesmo que estas perdas sejam menores, a possibilidade delas estarem relacionadas com a medida dependente ou uma variável independente, podem levar a resultados enviesados.

Neste trabalho, utilizou-se uma base de dados já registrados, ou seja, uma coorte retrospectiva. Um primeiro passo na análise de sobrevida em estudos de coorte que utilizam dados já registrados é a atualização do status do paciente. Na atualização efetuada pela autora, neste estudo, os perdidos de seguimento antes do trabalho de recuperação de informações representavam 23,5% dos indivíduos. Após a consulta nos prontuários dos pacientes, apesar de muitos terem retornado ao serviço e outros terem ido a óbito durante o período, ao final esta proporção não sofreu grande modificação, passando para 22,9%.

Na busca dos perdidos de seguimento, foi constatado nos prontuários que, na maior parte, já havia registros de tentativas de contato através de cartas para o endereço do paciente e, na falha deste tipo de contato, eram enviadas cartas aos órgãos de segurança da região de moradia. Toda a

informação possível, além dos registros de mortalidade que são usualmente utilizados pelo serviço, é utilizada para a atualização do status dos pacientes.

Dos 522 pacientes perdidos de seguimento, 267 (51%) já possuíam este status no banco de dados original. Analisando estes pacientes, vê-se que quase a metade deles (120 ; 44%) já deixaram de comparecer ao serviço ou não foram localizados no ano de 1994. Mesmo com a existência de uma equipe de funcionários para efetuar a busca ativa de pacientes (através de telefone, telegramas e cartas) que não comparecem ao serviço, uma parte dos pacientes é perdida de seguimento. Isso mostra que, apesar de grande esforço na recuperação de informações sobre os pacientes, o pesquisador deve estar ciente de que poderá ter muita dificuldade em chegar a percentuais de perda aceitáveis.

O planejamento do estudo com a definição dos procedimentos para a busca dos pacientes e do encerramento do seguimento foram citados por CARVALHO e col. (2005) como medidas a serem implementadas em qualquer estudo para diminuir a perda de pacientes.

WEICHUNG (2002) recomendou os seguintes passos para diminuir a perda de seguimento em estudos de coorte ou ensaios clínicos: coletar as razões das perdas de seguimento, conduzir as análises utilizando diferentes cenários e discutir as diferenças entre elas, dar atenção às possíveis causas de perda de seguimento no planejamento e durante o monitoramento do estudo, coletar o máximo de informações possíveis do período pós-perda de seguimento quando existir mais de um tipo de falha considerada, atribuir a perda de seguimento em si um resultado importante quando ela for grande. No caso deste estudo, em que os pacientes foram analisados retrospectivamente, só foi possível realizar algumas destas recomendações, tais como o estudo dos diferentes cenários e análise das perdas entre os grupos.

Em que pese sua importância, nem sempre a informação sobre a quantidade de observações perdidas de seguimento é fornecida nos estudos

de sobrevida. ALTMAN e col. (1991) realizaram revisão dos artigos publicados em cinco revistas especializadas em câncer e só em 34 dos 132 (26%) artigos revisados foram mencionadas as quantidades de perdas, 12 informaram não ter tido perdas e dos 22 restantes, somente 10 informaram como elas foram tratadas.

Por outro lado, como seria de se esperar em uma coorte de pacientes de câncer, a taxa de mortalidade dos pacientes foi maior após os procedimentos de atualização das informações do banco de dados, passando de 27,6% para 43,4% de óbitos. Estes resultados foram semelhantes aos encontrados por FAREWELL e col. (2003) na análise do efeito da perda de seguimento de pacientes com lupus eritematoso, onde o percentual de óbitos foi maior após a atualização do status, diminuindo a probabilidade de sobrevida.

A atualização do banco de dados é um passo importante em análise de sobrevida, pois neste trabalho, se fossem utilizadas as informações do banco de dados original, os resultados estariam errados. Dos 1117 pacientes considerados vivos no banco original, 221 haviam falecido, e 255 deixaram de comparecer ao serviço e, dentre os 536 pacientes considerados perdidos de seguimento, 160 haviam morrido, 109 voltaram ao serviço e continuavam vivos e 267 permaneceram sem comparecer ao serviço nos cinco anos de seguimento. Então apesar da proporção das perdas não ter sofrido diferenças, a busca e atualização dos pacientes geraram uma grande mudança na base de dados, elevando a proporção de óbitos de 27,6% para 43,5%.

Tu e col. (1999) avaliaram algumas possíveis estratégias de análise das perdas de seguimento para os pacientes de AIDS registrados no sistema de vigilância dos Estados Unidos. Uma delas foi considerar os casos sem óbito comunicado como vivo ao final do estudo (censura), apontando para a possível superestimação dos parâmetros ao utilizar esta estratégia. Outra tática foi a exclusão destes casos sem registro de seguimento e a posterior análise só dos casos em que o status era conhecido. Os autores

referem que os resultados com a exclusão dos pacientes, apesar de ser também sujeito a viéses, apresentaram probabilidades de sobrevida mais próximas da realidade, estimadas através de alguns ensaios clínicos.

HENNEKENS e BURING (1987) indicam como medida indireta utilizada para medir viés introduzido pela perda de seguimento (em estudos de coorte) é fazer as análises assumindo a situação extrema, ou seja comparando os resultados em que todos os que foram perdidos sofreram a falha com o cenário em que todos os perdidos não sofreram a falha, assim teríamos uma idéia da amplitude das estimativas dos parâmetros, pois neste intervalo o verdadeiro valor se situaria. Os autores observam que se as perdas forem muito grandes, este intervalo também o será. Seguindo este raciocínio, para o presente estudo, os verdadeiros valores das probabilidades de sobrevida poderiam estar entre 70 (com a análise sem os perdidos de seguimento) e 77% (considerando todos vivos ao final do tempo de seguimento) para um ano, 51 e 62% para três anos e 44 a 57% para cinco anos.

A comparação das características dos pacientes que deixaram de comparecer ao serviço (perdidos de seguimento) com o grupo dos pacientes não perdidos, isto é, que sofreram a falha ou estavam vivos pelo término do período de observação, revelou que as maiores proporções de perdas ocorreram em pacientes com maior escolaridade, categoria de admissão por convênio ou particular, menos graves (estádios I ou II) e nos pacientes com idade entre 15 e 39 anos e com mais de 64 anos. Com isso, pode-se levantar a hipótese de que muitos desses pacientes possam estar vivos. Caso isso seja verdadeiro, esta é uma perda informativa a favor dos vivos e as análises da base de dados padrão podem estar sub-estimadas.

A avaliação das perdas anuais do período de seguimento revelou que, quando os pacientes perdidos de observação são retirados a cada ano, as probabilidades de sobrevida diminuem, como esperado. A retirada de todos os pacientes perdidos de seguimento acarreta uma queda de 7% na probabilidade de sobrevida com cinco anos de seguimento. Quando esta

avaliação é feita levando-se em conta as variáveis de estudo, nota-se que, na variável sexo, as diferenças entre os valores da probabilidade do banco de dados com todas as perdas e os bancos sem as perdas são maiores para o sexo masculino; já na variável cor/raça, estas diferenças não são acentuadas. Na categoria de admissão, há maior diferença entre as probabilidades dos pacientes admitidos pelo SUS. A variável estágio também apresentou maiores diferenças entre as categorias, sendo que os pacientes mais graves tiveram as maiores diferenças. Estes fatos ocorrem, pois em algumas variáveis, a proporção de perda foi maior para algumas categorias. Os pacientes com mais de 64 anos também apresentaram maior diferença nas probabilidades de sobrevida sem as perdas anuais quando comparadas com a base de dados padrão.

Neste estudo, quando as perdas de seguimento foram transformadas em vivos ao final do período de 5 anos de seguimento, a probabilidade de sobrevida passou de 0,75 para 0,77 no primeiro ano, de 0,58 para 0,62 no terceiro ano e de 0,50 para 0,57 no quinto ano. Portanto, houve um aumento de 14% na estimativa da sobrevida em cinco anos. Com relação à variável sexo, os homens tiveram 20% e as mulheres 10% de aumento na probabilidade de sobrevida em cinco anos. Os pacientes com estágio I ou II tiveram aumento de 4%, enquanto os com estágio III ou IV tiveram 30% de aumento na probabilidade de sobrevida em cinco anos.

Quando as perdas foram transformadas em vivos ao final do seguimento de 5 anos, as HR's sofreram quedas. Na variável sexo a queda foi de 3,5% e na faixa etária ela variou de 5% (para os menores de 15 anos) a 14% (para os maiores de 64 anos). Na variável estágio, o HR passou de 5,41 para 4,88, uma queda de 9,8%. A transformação dos perdidos em vivos gera uma diminuição da força de morbidade com relação às variáveis de estudo.

As técnicas de análise de sobrevida têm como suposição básica que os motivos que levaram à perda de seguimento não tenham relação com o tempo de sobrevida e a comparação dessas perdas de seguimento segundo

algumas características dos pacientes é uma das formas de se avaliar o tipo de censura. Outra medida indicada por LEUNG e col. (1997) para se avaliar as perdas de seguimento, embora advertindo aos pesquisadores a possibilidade de aumento nos gastos com a pesquisa, seria proceder a busca ativa de uma parte dos pacientes perdidos de seguimento após o término do estudo e comparar suas características e as sobrevidas residuais dos dois grupos. Este procedimento daria uma indicação se a utilização das perdas de seguimento estaria super ou sub estimando as probabilidades de sobrevida.

LEE e GO (1997) afirmam que uma importante suposição do método Kaplan-Meier é que a probabilidade de uma observação ser censurada é independente de seu tempo de seguimento, ou, em outras palavras, que a razão de uma observação ser censurada não deve estar relacionada com a falha estudada. ANDERSEN e col. (1993) afirmam que em análise de sobrevida, a perda de seguimento deve ser independente da sobrevida e que se a censura for independente em qualquer tempo t , a sobrevida não será alterada pela censura.

CARVALHO e col. (2005) citam que se o indivíduo foi perdido em decorrência de causa associada ao evento estudado, como, por exemplo, a piora do paciente ou se o óbito não for conhecido por falha de acompanhamento, esta pode ser denominada censura informativa. SZKLO & NIETO (2004) apontam que a imigração e o óbito por outras causas, podem representar possíveis violações da independência entre a censura e a probabilidade de sobrevida. Os autores também consideram que a censura pelo término do estudo tem poucas chances de representar algum viés nas análises de sobrevida.

Neste estudo, a avaliação do tipo de perda foi realizada quando uma parte dos óbitos foi recodificado como censura, simulando uma perda informativa. Na base de dados, a transformação de 10% dos óbitos em censura acarretaram aumentos de 4%, 6% e 10% nas probabilidades com um, três e cinco anos de seguimento. Com 15% dos óbitos transformados

em censura, estas diferenças aumentam para 6%, 12% e 17% com um, três e cinco anos de seguimento. Com 50% de perda as diferenças aumentam para 20%, 39% e 53% com um, três e cinco anos de seguimento.

Analisando os mesmos bancos de dados segundo as variáveis de estudo, nota-se que as diferenças entre as probabilidades de sobrevida acumuladas calculadas na base de dados padrão com as perdas informativas são semelhantes, independente da variável de estudo. Também foi observado que, as categorias com menor probabilidades de sobrevida sofrem diferenças percentuais maiores se comparadas às de maior probabilidade. Por exemplo, na variável sexo, a simulação de 10% de perda acarretou acréscimo de 1% para o sexo feminino e 8% para o sexo masculino. Na variável categoria de admissão, as diferenças entre os pacientes admitidos pelo SUS e por convênio ou particulares foram semelhantes.

Portanto, a possibilidade das perdas de seguimento serem do tipo informativas deve ser sempre analisada, visto que, se elas forem superiores a 10%, o uso do estimador de Kaplan-Meier pode acarretar uma super ou sub estimação das probabilidades de sobrevida acumuladas e estas diferenças podem se acentuar dependendo do tempo de seguimento e das variáveis de estudo.

No entanto, nas HR's, quando as perdas simuladas foram informativas nos óbitos, as diferenças máximas alcançadas ficaram em torno de 5% (15% de perda) até 25% (até 50% de perda). Por exemplo, a variável estágio apresentou aumento de 3% com 50% de perda informativa nos óbitos, enquanto que na variável escolaridade, este procedimento acarretou aumento de 8% na força de mortalidade, contudo, nenhuma destas mudanças apresentaram uma tendência significativa com relação aos percentuais de perda. Estes resultados indicam que elas são menos afetadas pelas perdas de seguimento, mesmo quando são informativas.

SIMON e WITTES (1986) recomendam que o máximo de perda aceitável seja de 15% dos pacientes elegíveis para um ensaio clínico. E que

a falta de seguimento por morte prematura, falta de informação ou violação do protocolo de mais de 20% dos pacientes refletem em problemas na seleção dos pacientes.

Quando as perdas foram não informativas, mesmo com a simulação de perda de 50%, a diferença máxima alcançada ficou em torno de 8%. Com isso, sugere-se que é imprescindível avaliar se as perdas são informativas ou não informativas nas estimativas das HR's e das probabilidades de sobrevida não simplesmente desconsiderar estes pacientes.

O mesmo aconteceu nas análises em todas as variáveis de estudo. Estes resultados indicam que quando a suposição para a utilização dos métodos de Kaplan-Meier e dos riscos proporcionais de Cox estão satisfeitas, independente da porcentagem de perdas, não há super ou sub estimação relevantes.

Sugere-se que, para ter uma idéia se a perda é informativa ou não que sejam comparadas algumas características demográficas e clínicas entre os perdidos de seguimento e os outros pacientes.

De maneira geral, a falta de independência dos tempos de censura com as falhas é um conceito simples de entender, entretanto a classificação do tipo de censura e a falta de testes disponíveis nos pacotes estatísticos, se constituem em séria dificuldade para o pesquisador que queira fazer análises de sobrevida de uma coorte de pacientes. Ao contrário das suposições de outras técnicas que podem ser facilmente verificadas. Por exemplo, no teste t de Student e na ANOVA há testes para igualdade de variância (Levéne) e de aderência à distribuição normal (Kolmogorov-Smirnov) disponíveis na maior parte dos pacotes estatísticos.

Concluindo, sabe-se que o percentual de perda de seguimento deve ser informado em todos os estudos para a avaliação dos possíveis vieses envolvidos na análise e a comparação das proporções de perdas entre as variáveis de estudo é um passo importante para a avaliação se essa perda é informativa ou não. Quando a comparação dos perdidos de seguimento e não perdidos não apresentar diferença entre as variáveis, pode-se

considerá-la como não informativa e, neste caso, no geral, as estimativas das probabilidades de sobrevida acumuladas realizadas pelo método Kaplan-Meier e das HR's pelo modelo dos riscos proporcionais de Cox podem não estar afetadas com perdas de seguimento de até 50%.

Caso a comparação das perdas indique a possibilidade da perda ser informativa, quanto maior sua proporção, maior a possibilidade de super ou sub estimação das probabilidades de sobrevida acumuladas, pelo método de Kaplan-Meier, aceitando-se um máximo de 15% de perdas de seguimento. Estas diferenças podem aumentar, principalmente após 3 anos.

Já para as estimativas das HR's, um percentual de 15% de perda acarreta diferenças nas estimativas ao redor de 5%, uma perda de seguimento de até 20% levou a aumento de até 20% nas estimativas dos parâmetros e perdas de seguimento de até 50% as maiores diferenças percentuais alcançadas nas HR's foram de 25%

6. CONCLUSÕES

6.1 O procedimento de atualização das informações sobre o status do paciente é um procedimento fundamental para análise de sobrevida.

6.2 Comparação da proporção de perdas de seguimento segundo as variáveis de estudo é um dos primeiros passos para análise do tipo de perda.

6.3 As probabilidades de sobrevida são diretamente afetadas pelo percentual de censura informativa.

6.4 Com mais de 15% de perda informativa e um período superior a 3 anos de seguimento, as probabilidades de sobrevida acumuladas estimadas pelo KM podem estar super ou sub estimadas em mais de 10%

6.5 As estimativas das HR's realizadas pelo modelo de riscos proporcionais de Cox são menos afetadas pela censura informativa

6.6 Caso a perda seja não informativa (aleatória) o método do estimador de Kaplan Meier e o estimador dos riscos proporcionais de Cox podem ser utilizados sem ressalvas

REFERÊNCIAS

1. Altman DG, De Stavola BL, Love SB, Stepniowska, Review of survival analyses published in cancer journals. **British Journal of Cancer**, 1995; 72:p.511-518.
2. Andersen PG, Borgan O, Gill RD, Keiding. **Statistical Models Based on Counting Processes**. Springer, 1991
3. Bustamente-Teixeira MT, Faerstein E, Latorre MRO. Técnicas de Análise de Sobrevida. **Cad. Saúde Pública**, 2002;18:p. 579-594.
4. Carvalho MS, Andreozzi VL, Codeço CT. Barbosa MTS, Shimakura. **Análise de sobrevida :Teoria e aplicações em saúde**. Rio de Janeiro. Editora Fiocruz, 2005.
5. Clark TG, Altman DG, Stavola BLD, Quantification of the completeness of follow-up. **The Lancet**, 2002; 359: p. 1309-1310.
6. Farewell VT, Lawless JF, Gladman DD, Urowitz MB. Tracing studies and analysis of the effect of loss to follow up on mortality estimation from patient registry data. **Appl. Statist.** 2003; 52. Part 4, p.455-6 .
7. Fleming TR, Lin DY, Survival Analysis in Clinical Trials: Past Developments and Future Directions. **Biometrics**, 2000; 56: p. 971-983.
8. Hennekens CH, Buring JE. **Epidemiology in Medicine**. Lippincot Williams & Wilkins,1987
9. Hosmer DW, Lemeshow S. **Applied Survival Analysis: regression modeling of time to event data**. Wiley series in probability and statistics, 1999.
10. Hougaard P. Fundamentals of Survival Data. **Biometrics**,1999;55: p.13-22.
11. Kleinbaum DG., **Survival Analysis: a self-learning text**. New York: Springer, 1995.
12. Lawless JF. **Statistical models and methods for lifetime data**. Wiley Series in probability and mathematical statistics,1982.
13. Lagakos SW, Willians JS. Models for censored survival analysis: A cone class of variable-sum models. **Biometrika** , 1978; 85:181-9.
14. Lagakos SW. General Right Censoring and Its Impact on the Analysis of Survival Data. **Biometrics**, 1979; 35: p. 139-156.

15. Latorrre MRDO. **Comparação entre Alguns Métodos Estatísticos em Análise de sobrevivência: Aplicação em uma Coorte de Pacientes com Câncer do Pênis**. São Paulo, 1996. [Tese de Doutorado – Faculdade de Saúde Pública da USP].
16. Latorrre MRDO. Análise de sobrevivência. In: Massad E, Menezes RX, Silveira PSP, Ortega NRS.et al. **Métodos Quantitativos em medicina**. São Paulo:Manole;2004. p.351.
17. Lee ET, Go OT. Survival Analysis in Public Health Research. **Annu. Rev. Public Health**, 1997; 18: p. 105-34.
18. Lee ET, Wei LJ, Amato DA. **Statistical methods for Survival data Analysis**. New York. Wiley 2nd edition, 1992.
19. Leung KM, Elashoff RM, Afifi AA. Censoring Issues in Survival Analysis. **Annu. Rev. Public Health**, 1997; 18: p. 83-104.
20. Lindsey JC & Ryan LM. Tutorial in biostatistics Methods for interval-censored data. **Statistics in Medicine**, 1998; 17: p. 219-238.
21. Link WA, A Model for Informative Censoring. **Journal of the American Statistical Association**, 1989; 84: p. 749-752.
22. Moher D, Schuls KF, Altman DG, The CONSORT statement: revised recommendations for improving the quality of reports of parallel-group randomized trials. **The Lancet**, 2001; 357: p. 1191-1194.
23. Simon R, Wittes RE. Methodologic guidelines for Reports of Clinical Trials. **Cancer**, 1986; 58: p. 212-4.
24. Szklo M & Nieto FJ .**Epidemiology: beyond the basics**. Jones and Bardett Publishers Inc.; 2004.
25. Toulomi G, Pocock SJ, Babiker AG, Darbyshire JH. Estimation and comparison of rates of change in longitudinal studies with informative drop-outs. **Statistics in Medicine**, 1999;18: p. 1215-1233.
26. Tu XM, Meng XL, Pagano M. The AIDS epidemic: estimating survival after AIDS diagnosis from surveillance data. **Journal of the American Statistical Association**, 1993; 88 p. 26-36.
27. Weichung JS. Problems in dealing with missing data and informative censoring in clinical trials Current **Controlled Trials in Cardiovascular Medicine**, 2002;3:4

ANEXO 1

FUNDAÇÃO ANTÔNIO PRUDENTE		REGISTRO HOSPITALAR DE CÂNCER - FICHA DE REGISTRO
I) IDENTIFICAÇÃO:		
01. Nº PRONTUÁRIO:	_____	
02. CATEGORIA:	11 SUS _____	
03. NOME:	_____	
04. DATA DE NASC.:	____/____/____	
05. IDADE:	_____ ANOS.	
06. SEXO:	11 MASC. 12 FEM. 19 IIGN _____	
07. RAÇA:	11 BRANCA 12 PRETA 13 PARDA 14 AMARELA 19 IIGN _____	
08. NATURALIDADE:	CIDADE: _____ ESTADO: _____	
	CEP: _____	
	PAÍS: _____	
09. RESIDÊNCIA - CIDADE:	_____ ESTADO: _____	
	CEP: _____	
	PAÍS: _____	
10. EST. CIVIL:	11 SOL. 12 CAS/UNIÃO LIVRE 13 SEP. 14 VIÚVO 19 IIGN _____	
11. ESCOLARIDADE:	10 ANalf. 11 SEMI ALF. 12 1º GRAU INCOMPLETO 13 1º GRAU COMPLETO 14 2º GRAU INCOMPL. 15 2º GRAU COMPLETO 16 SUPER. INCOMPL. 17 SUP. COMPL. 19 IIGN _____	
12. ATIV. PROFISSIONAL:	_____	
II) INFORMAÇÕES SOBRE A DOENÇA:		
13. DATA ADMISSÃO:	____/____/____	
14. DEPARTAMENTO:	_____ Nº: _____	
15. ORIGEM DO PACIENTE:	11 CLÍNICA 12 REDE PÚBLICA _____ 13 POR CONTA PRÓPRIA 19 IIGN _____	
16. DIAGNÓSTICO PRÉVIO:	10 NÃO 11 SIM 19 IIGN _____	
17. DATA DO DIAGNÓSTICO:	____/____/____	
18. TRATAMENTO PRÉVIO:	10 NÃO 11 CIR. 12 RXT 13 QT 14 CIR+RXT 15 CIR+QT 16 RXT+QT 17 CIR+RXT+QT 18 HORMONIOTERAPIA 19 OUTROS _____	
19. BASE MAIS IMPORTANTE PARA O DIAGNÓSTICO:	11 APENAS CLÍNICO _____ 12 IRX. ENDOSC. ETC 13 CIR. EXPLORATÓRIA 14 TESTES LABORAT. 15 CITO/HEMATOLOGIA 16 HISTOLOGIA/METÁ 17 HISTOL. TU PRIMÁRIO 18 INECRÓPSIA 19 IIGN _____	
20. TOPOGRAFIA DO TU PRIMÁRIO:	_____	
	T CID-O: _____	
21. MORFOLOGIA DO TU PRIMÁRIO:	_____	
	M CID-O: _____	
22. TOPOGRAFIA DA METÁSTASE:	10 NÃO 11 FÍGADO 12 PULMÃO 13 OSSO 14 CÉREBRO 15 PELE 16 GANGLIONAR (_____)	

24. DOENÇA NÃO NEOPLÁSICA: _____
CID: _____ ☐ ☐ ☐ ☐

25. TUMORES PRIMÁRIOS MÚLTIPLOS: 10 INÃO 11 ISIM 12 DUVIDOSO 19 IIGN ☐

26. EXTENSÃO CLÍNICA DA DOENÇA ANTES DO TRATAMENTO:
A) ESTADIAMENTO T: _____ N: _____ M: _____ ☐ ☐ ☐
B) ESTADIAMENTO CLÍNICO: _____ ☐

27. DATA DO INÍCIO DO TRATAMENTO: ____/____/____, ____/____/____

28. TRATAMENTO REALIZADO NA INSTITUIÇÃO: 10 INÃO 11 CIR 12 RXT ☐
13 IQT 14 CIR+RXT 15 CIR+QT 16 RXT+QT 17 CIR+RXT+QT
18 HORMONIO TERAPIA 19 OUTROS: (_____)

29. ESTADO DO DOENTE AO FINAL DO TRATAMENTO: _____ ☐
11 SEM EVIDÊNCIA DE DOENÇA 12 DOENÇA EM ATIVIDADE
13 EM REGRESSÃO 14 INALTERADO 15 EM TRATAMENTO
16 ÓBITO ANTES DO TÉRMINO DO TRAT//O (DT. ____/____/____) ____/____/____
CAUSA MORTIS: (_____) 19 IIGN

III) TUMORES MÚLTIPLOS:

30. SEGUNDO TUMOR PRIMÁRIO:
A) TOPOGRAFIA: _____
T CID-O: _____ ☐ ☐ ☐ ☐
B) MORFOLOGIA: _____
M CID-O: _____ ☐ ☐ ☐ ☐
C) DATA DO DIAGNÓSTICO: ____/____/____, ____/____/____
D) TNM CLÍNICO: T: _____ N: _____ M: _____ ☐ ☐ ☐ ☐
E) ESTADIAMENTO CLÍNICO: _____ ☐

31. TERCEIRO TUMOR PRIMÁRIO:
A) TOPOGRAFIA: _____
T CID-O: _____ ☐ ☐ ☐ ☐
B) MORFOLOGIA: _____
M CID-O: _____ ☐ ☐ ☐ ☐
C) DATA DO DIAGNÓSTICO: ____/____/____, ____/____/____
D) TNM CLÍNICO: T: _____ N: _____ M: _____ ☐ ☐ ☐ ☐
E) ESTADIAMENTO CLÍNICO: _____ ☐

32. RECIDIVA: 10 INÃO 11 LOCAL 12 GANGLIONAR ☐
13 DISTÂNCIA (_____) 19 IIGN

33. DATA DA 1ª RECIDIVA: ____/____/____, ____/____/____

34. TRATAMENTO DA RECIDIVA: 10 INÃO 11 ISIM ☐

35. TEVE MAIS DE TRÊS TUMORES PRIMÁRIOS ? 10 INÃO 11 ISIM ☐

IV. SEGUIMENTO

DATA	OBSERVAÇÃO	CÓDIGO *
____/____/____	_____	_____

[illegible]

ANEXO 2

São Paulo, 30 de outubro de 2002

À

Dra. Márcia Moreira Holcman

Ref. Comunicação da Comissão de Ética em Pesquisa

Projeto: "Comparação do Impacto das Perdas de Seguimento nas Análises Feitas pelo Estimador Produto Limite de Kaplan Meier e Modelo de Riscos Proporcionais de Cox". (CEP – n.º 449/02)

Prezado Dra. Márcia,

Seu projeto de pesquisa, acima mencionado, foi apreciado pela Comissão de Ética em Pesquisa (CEP) do Hospital do Câncer em sua última reunião de 29/10/2002. Os membros desta comissão aprovaram a realização deste estudo. Informações a respeito do andamento do referido projeto deverão ser encaminhadas à secretaria do CEP dentro de 12 meses.

Atenciosamente,

Dr. Agnaldo Anelli
Presidente da Comissão de Ética em Pesquisa

C/c: Dra. Maria do Rosário Dias de Oliveira Latorre

Dr. Luiz Paulo Kowalski

Dra. Karina de Cássia Braga Ribeiro

Comissão de Ética em Pesquisa - CEP
Rua Prof. Antonio Prudente, 211 – Liberdade – São Paulo – SP
CEP – 01509-900 – Tel. (0**11) 3272-5000 ramal 1117

ANEXO 3

Universidade de São Paulo

Faculdade de Saúde Pública

COMITÊ DE ÉTICA - COEP

Av. Dr. Arnaldo, 715 - CEP 01246-904 - São Paulo - Brasil

Telefones: (55-11) 3066-7734 - fone/fax (55-11) 3064-7314 - e-mail: mdgracas@usp.br

Of.COEP/269/02

13 de novembro de 2002

Pelo presente, informo que o Comitê de Ética em Pesquisa da Faculdade de Saúde Pública da Universidade de São Paulo-COEP, **analisou e aprovou com recomendações**, em sua 9.ª/02 Sessão Ordinária, realizada em 12.11.02, de acordo com os requisitos da Resolução 196/96, o Protocolo de Pesquisa n.º 870, intitulado: "COMPARAÇÃO DO IMPOSTO DAS PERDAS DE SEGUIMENTO NAS ANÁLISES FEITAS PELO ESTIMADOR PRODUTO LIMITE DE KAPLAN MEIER E MODELO DE RISCOS PROPORCIONAIS DE COX", apresentado pela pesquisadora Márcia Moreira Holcman.

Considera que a pesquisadora deva elaborar um Termo de Compromisso para a instituição participante, assegurando o uso previsto dos dados e a garantia de manutenção estrita da confidencialidade.

Atenciosamente,

Paulo Antonio de Carvalho Fortes
Professor Associado
Coordenador do Comitê de Ética em Pesquisa da FSP-COEP

24. DOENÇA NÃO NEOPLÁSICA: _____
CID: _____ ☐ ☐ ☐ ☐

25. TUMORES PRIMÁRIOS MÚLTIPLOS: 10 NÃO 11 SIM 12 DUVIDOSO 19 IGN ☐

26. EXTENSÃO CLÍNICA DA DOENÇA ANTES DO TRATAMENTO:
A) ESTADIAMENTO T: _____ N: _____ M: _____ ☐ ☐
B) ESTADIAMENTO CLÍNICO: _____ ☐

27. DATA DO INÍCIO DO TRATAMENTO: ____/____/____

28. TRATAMENTO REALIZADO NA INSTITUIÇÃO: 10 NÃO 11 CIR. 12 RXT ☐
13 QT 14 CIR+RXT 15 CIR+QT 16 RXT+QT 17 CIR+RXT+QT
18 HORMONIO TERAPIA 19 OUTROS: (_____)

29. ESTADO DO DOENTE AO FINAL DO TRATAMENTO: _____ ☐
11 SEM EVIDÊNCIA DE DOENÇA 12 DOENÇA EM ATIVIDADE
13 EM REGRESSÃO 14 INALTERADO 15 EM TRATAMENTO
16 ÓBITO ANTES DO TÉRMINO DO TRAT//O (DT. ____/____/____) ____/____/____
CAUSA MORTIS: (_____) 19 IGN

III) TUMORES MÚLTIPLOS:

30. SEGUNDO TUMOR PRIMÁRIO:
A) TOPOGRAFIA: _____
T CID-O: _____ ☐ ☐ ☐ ☐
B) MORFOLOGIA: _____
M CID-O: _____ ☐ ☐ ☐ ☐
C) DATA DO DIAGNÓSTICO: ____/____/____
D) TNM CLÍNICO: T: _____ N: _____ M: _____ ☐ ☐ ☐ ☐
E) ESTADIAMENTO CLÍNICO: _____ ☐ ☐

31. TERCEIRO TUMOR PRIMÁRIO:
A) TOPOGRAFIA: _____
T CID-O: _____ ☐ ☐ ☐ ☐
B) MORFOLOGIA: _____
M CID-O: _____ ☐ ☐ ☐ ☐
C) DATA DO DIAGNÓSTICO: ____/____/____
D) TNM CLÍNICO: T: _____ N: _____ M: _____ ☐ ☐ ☐ ☐
E) ESTADIAMENTO CLÍNICO: _____ ☐ ☐

32. RECIDIVA: 10 NÃO 11 LOCAL 12 GANGLIONAR ☐
13 DISTÂNCIA (_____) 19 IGN

33. DATA DA 1ª RECIDIVA: ____/____/____

34. TRATAMENTO DA RECIDIVA: 10 NÃO 11 SIM ☐

35. TEVE MAIS DE TRÊS TUMORES PRIMÁRIOS ? 10 NÃO 11 SIM ☐

IV. SEGUIMENTO

DATA	OBSERVAÇÃO	CÓDIGO *
____/____/____	_____	_____

<i>Banco</i>	<i>Ano</i>		
	1 ano	3 ano	5 ano
0	0,70	0,51	0,44
5	0,71	0,53	0,47
10	0,73	0,54	0,48
15	0,74	0,58	0,52
20	0,75	0,59	0,53
25	0,77	0,60	0,55
30	0,78	0,63	0,57
35	0,80	0,66	0,59
40	0,81	0,67	0,62
45	0,82	0,69	0,64
50	0,84	0,71	0,68

Anexo 4 – Probabilidade de sobrevida acumulada do total de pacientes com perdas geradas nos óbitos até o primeiro, terceiro e quinto ano de seguimento, RHC, 1994.

Anexo 5 – Probabilidade de sobrevida acumulada do total de pacientes com perdas aleatórias até o primeiro, terceiro e quinto ano de seguimento, RHC, 1994.

<i>Banco</i>	<i>Ano</i>		
	1 ano	3 ano	5 ano
0	0,75	0,58	0,50
5	0,76	0,58	0,50
10	0,75	0,58	0,50
15	0,75	0,57	0,50
20	0,75	0,58	0,50
25	0,76	0,57	0,51
30	0,74	0,56	0,49
35	0,77	0,58	0,50
40	0,76	0,58	0,51
45	0,74	0,57	0,50
50	0,75	0,59	0,51

Anexo 6– Probabilidade de sobrevida acumulada do total de pacientes com perdas geradas nos óbitos segundo sexo até o primeiro, terceiro e quinto ano de seguimento, RHC, 1994.

Banco	Sexo	Ano			Teste Log Rank (p)
		1 ano	3 ano	5 ano	
Total	feminino	0,78	0,61	0,53	p<0,0001
	masculino	0,61	0,41	0,33	
O5	feminino	0,79	0,62	0,55	p<0,0001
	masculino	0,62	0,43	0,37	
O10	feminino	0,79	0,62	0,56	p<0,0001
	masculino	0,65	0,45	0,39	
O15	feminino	0,82	0,67	0,61	p<0,0001
	masculino	0,65	0,47	0,41	
O20	feminino	0,82	0,68	0,62	p<0,0001
	masculino	0,67	0,49	0,43	
O25	feminino	0,84	0,69	0,64	p<0,0001
	masculino	0,68	0,49	0,44	
O30	feminino	0,84	0,70	0,65	p<0,0001
	masculino	0,70	0,54	0,48	
O35	feminino	0,86	0,74	0,68	p<0,0001
	masculino	0,73	0,55	0,49	
O40	feminino	0,87	0,75	0,69	p<0,0001
	masculino	0,73	0,58	0,52	
O45	feminino	0,87	0,76	0,71	p<0,0001
	masculino	0,77	0,61	0,55	
O50	feminino	0,88	0,78	0,75	p<0,0001
	masculino	0,78	0,62	0,58	

Anexo 7– Probabilidade de sobrevida acumulada do total de pacientes com perdas aleatórias segundo sexo até o primeiro, terceiro e quinto ano de seguimento, RHC, 1994.

Banco	Sexo	Ano			Teste Log Rank (p)
		1 ano	3 ano	5 ano	
Total	feminino	0,82	0,66	0,58	p<0,01
	masculino	0,67	0,48	0,40	
AL5	feminino	0,82	0,66	0,58	p<0,01
	masculino	0,68	0,49	0,41	
AL10	feminino	0,82	0,66	0,59	p<0,01
	masculino	0,67	0,49	0,40	
AL15	feminino	0,82	0,66	0,58	p<0,01
	masculino	0,67	0,48	0,41	
AL20	feminino	0,82	0,67	0,60	p<0,01
	masculino	0,67	0,47	0,39	
AL25	feminino	0,81	0,65	0,59	p<0,01
	masculino	0,69	0,48	0,42	
AL30	feminino	0,82	0,66	0,59	p<0,01
	masculino	0,66	0,46	0,38	
AL35	feminino	0,82	0,66	0,57	p<0,01
	masculino	0,70	0,48	0,41	
AL40	feminino	0,84	0,66	0,60	p<0,01
	masculino	0,68	0,48	0,40	
AL45	feminino	0,81	0,63	0,57	p<0,01
	masculino	0,67	0,50	0,41	
AL50	feminino	0,81	0,67	0,59	p<0,01
	masculino	0,69	0,48	0,41	

Anexo 8— Probabilidade de sobrevida acumulada do total de pacientes com perdas geradas nos óbitos segundo idade até o primeiro, terceiro e quinto ano de seguimento, RHC, 1994.

Banco	Idade	Ano			Teste Log Rank (p)
		1 ano	3 ano	5 ano	
Total	0-14	0,77	0,63	0,6	p<0,01
	15-39	0,8	0,65	0,58	
	40-64	0,73	0,54	0,46	
	65 e +	0,58	0,38	0,3	
O5	0-14	0,77	0,63	0,61	p<0,01
	15-39	0,81	0,65	0,6	
	40-64	0,74	0,55	0,47	
	65 e +	0,6	0,4	0,32	
O10	0-14	0,78	0,64	0,62	p<0,01
	15-39	0,82	0,68	0,62	
	40-64	0,76	0,57	0,51	
	65 e +	0,61	0,41	0,34	
O15	0-14	0,8	0,67	0,65	p<0,01
	15-39	0,81	0,67	0,61	
	40-64	0,77	0,6	0,54	
	65 e +	0,64	0,45	0,39	
O20	0-14	0,8	0,68	0,65	p<0,01
	15-39	0,83	0,7	0,64	
	40-64	0,78	0,62	0,53	
	65 e +	0,65	0,45	0,39	
O25	0-14	0,82	0,69	0,66	p<0,01
	15-39	0,85	0,72	0,67	
	40-64	0,79	0,62	0,56	
	65 e +	0,67	0,49	0,43	
O30	0-14	0,85	0,75	0,74	p<0,01
	15-39	0,86	0,73	0,68	
	40-64	0,8	0,65	0,59	
	65 e +	0,68	0,51	0,45	
O35	0-14	0,88	0,77	0,73	p<0,01
	15-39	0,86	0,75	0,7	
	40-64	0,81	0,66	0,6	
	65 e +	0,73	0,55	0,48	
O40	0-14	0,87	0,79	0,76	p<0,01
	15-39	0,86	0,78	0,73	
	40-64	0,84	0,68	0,62	
	65 e +	0,72	0,57	0,49	
O45	0-14	0,88	0,77	0,74	p<0,01
	15-39	0,87	0,79	0,74	
	40-64	0,84	0,7	0,64	
	65 e +	0,76	0,62	0,55	
O50	0-14	0,91	0,83	0,83	p<0,01
	15-39	0,88	0,79	0,77	
	40-64	0,86	0,73	0,68	
	65 e +	0,75	0,61	0,56	

Anexo 9– Probabilidade de sobrevida acumulada do total de pacientes com perdas aleatórias segundo idade até o primeiro, terceiro e quinto ano de seguimento, RHC, 1994.

Banco	Idade	Ano			Teste Log Rank (p)
		1 ano	3 ano	5 ano	
Total	0-14	0,8	0,87	0,64	p<0,01
	15-39	0,84	0,71	0,64	
	40-64	0,77	0,59	0,51	
	65 e +	0,66	0,46	0,38	
al5	0-14	0,77	0,63	0,61	p<0,01
	15-39	0,79	0,64	0,59	
	40-64	0,73	0,54	0,46	
	65 e +	0,58	0,38	0,31	
al10	0-14	0,8	0,66	0,63	p<0,01
	15-39	0,84	0,71	0,64	
	40-64	0,77	0,6	0,51	
	65 e +	0,67	0,57	0,38	
al15	0-14	0,8	0,68	0,65	p<0,01
	15-39	0,71	0,66	0,77	
	40-64	0,59	0,5	0,65	
	65 e +	0,65	0,55	0,37	
al20	0-14	0,79	0,66	0,62	p<0,01
	15-39	0,71	0,64	0,77	
	40-64	0,6	0,52	0,66	
	65 e +	0,66	0,46	0,37	
al25	0-14	0,8	0,66	0,66	p<0,01
	15-39	0,84	0,64	0,77	
	40-64	0,6	0,52	0,67	
	65 e +	0,67	0,47	0,36	
al30	0-14	0,76	0,62	0,59	p<0,01
	15-39	0,84	0,71	0,67	
	40-64	0,76	0,57	0,49	
	65 e +	0,67	0,46	0,37	
al35	0-14	0,79	0,62	0,59	p<0,01
	15-39	0,84	0,7	0,64	
	40-64	0,8	0,59	0,5	
	65 e +	0,67	0,46	0,4	
al40	0-14	0,8	0,68	0,67	p<0,01
	15-39	0,83	0,7	0,63	
	40-64	0,76	0,59	0,52	
	65 e +	0,69	0,46	0,37	
al45	0-14	0,75	0,63	0,59	p<0,01
	15-39	0,83	0,67	0,62	
	40-64	0,77	0,59	0,5	
	65 e +	0,66	0,47	0,4	
al50	0-14	0,75	0,63	0,59	p<0,01
	15-39	0,83	0,67	0,62	
	40-64	0,77	0,59	0,5	
	65 e +	0,66	0,47	0,4	

Anexo 10– Probabilidade de sobrevida acumulada do total de pacientes com perdas geradas nos óbitos segundo cor/raça até o primeiro, terceiro e quinto ano de seguimento, RHC, 1994.

Banco	Cor/raça	Ano			Teste Log Rank (p)
		1 ano	3 ano	5 ano	
Total	branca ou amarela	0,70	0,52	0,45	0,40
	preta ou parda	0,69	0,48	0,41	
O5	branca ou amarela	0,71	0,54	0,47	0,40
	preta ou parda	0,70	0,50	0,43	
O10	branca ou amarela	0,73	0,55	0,49	0,50
	preta ou parda	0,72	0,50	0,44	
O15	branca ou amarela	0,74	0,58	0,52	0,30
	preta ou parda	0,74	0,53	0,47	
O20	branca ou amarela	0,75	0,60	0,54	0,50
	preta ou parda	0,74	0,55	0,49	
O25	branca ou amarela	0,77	0,61	0,55	0,50
	preta ou parda	0,76	0,57	0,52	
O30	branca ou amarela	0,78	0,63	0,58	0,10
	preta ou parda	0,77	0,60	0,54	
O35	branca ou amarela	0,80	0,66	0,60	0,70
	preta ou parda	0,81	0,64	0,56	
O40	branca ou amarela	0,81	0,68	0,62	0,20
	preta ou parda	0,81	0,63	0,56	
O45	branca ou amarela	0,82	0,70	0,64	0,80
	preta ou parda	0,83	0,67	0,62	
O50	branca ou amarela	0,83	0,72	0,68	0,90
	preta ou parda	0,81	0,66	0,61	

Anexo 11– Probabilidade de sobrevida acumulada do total de pacientes com perdas aleatórias segundo sexo até o primeiro, terceiro e quinto ano de seguimento, RHC, 1994.

Banco	Cor/raça	Ano			Teste Log Rank (p)
		1 ano	3 ano	5 ano	
Total	branca ou amarela	0,75	0,58	0,51	0,30
	preta ou parda	0,74	0,54	0,47	
AL5	branca ou amarela	0,75	0,59	0,51	0,30
	preta ou parda	0,76	0,54	0,46	
AL10	branca ou amarela	0,75	0,59	0,51	0,21
	preta ou parda	0,74	0,55	0,46	
AL15	branca ou amarela	0,75	0,58	0,51	0,23
	preta ou parda	0,74	0,53	0,46	
AL20	branca ou amarela	0,75	0,58	0,51	0,36
	preta ou parda	0,74	0,55	0,47	
AL25	branca ou amarela	0,76	0,58	0,52	0,30
	preta ou parda	0,75	0,52	0,47	
AL30	branca ou amarela	0,74	0,56	0,49	0,87
	preta ou parda	0,76	0,56	0,47	
AL35	branca ou amarela	0,77	0,58	0,50	0,34
	preta ou parda	0,74	0,54	0,47	
AL40	branca ou amarela	0,77	0,58	0,51	0,27
	preta ou parda	0,74	0,54	0,47	
AL45	branca ou amarela	0,75	0,59	0,51	0,01
	preta ou parda	0,69	0,46	0,41	
AL50	branca ou amarela	0,76	0,59	0,52	0,71
	preta ou parda	0,73	0,57	0,49	

Anexo 12– Probabilidade de sobrevivência acumulada do total de pacientes com perdas geradas nos óbitos segundo escolaridade até o primeiro, terceiro e quinto ano de seguimento. RHC. 1994.

Banco	Escolaridade	Ano			Teste Log Rank (p)
		1 ano	3 ano	5 ano	
Total	fund ou + até fund,	0,74	0,59	0,52	p<0,01
		0,69	0,49	0,41	
O5	fund ou + até fund,	0,75	0,60	0,53	p<0,01
		0,69	0,49	0,42	
O10	fund ou + até fund,	0,76	0,60	0,54	p<0,01
		0,70	0,51	0,44	
O15	fund ou + até fund,	0,77	0,61	0,57	p<0,01
		0,71	0,53	0,45	
O20	fund ou + até fund,	0,77	0,63	0,57	p<0,01
		0,72	0,55	0,47	
O25	fund ou + até fund,	0,78	0,65	0,60	p<0,01
		0,72	0,54	0,48	
O30	fund ou + até fund,	0,78	0,66	0,59	p<0,01
		0,74	0,57	0,50	
O35	fund ou + até fund,	0,79	0,68	0,64	p<0,01
		0,74	0,58	0,53	
O40	fund ou + até fund,	0,82	0,69	0,63	p<0,01
		0,77	0,61	0,54	
O45	fund ou + até fund,	0,81	0,71	0,66	p<0,01
		0,76	0,61	0,55	
O50	fund ou + até fund,	0,82	0,69	0,65	0,07
		0,78	0,66	0,60	

Anexo 13— Probabilidade de sobrevida acumulada do total de pacientes com perdas aleatórias segundo escolaridade até o primeiro, terceiro e quinto ano de seguimento, RHC, 1994.

Banco	Escolaridade	Ano			Teste Log Rank (p)
		1 ano	3 ano	5 ano	
Total	fund ou +	0,80	0,65	0,59	p<0,01
	até fund,	0,73	0,55	0,47	
AL5	fund ou +	0,80	0,65	0,59	p<0,01
	até fund,	0,74	0,55	0,47	
AL10	fund ou +	0,79	0,66	0,59	p<0,01
	até fund,	0,74	0,55	0,47	
AL15	fund ou +	0,79	0,65	0,59	0,01
	até fund,	0,73	0,55	0,47	
AL20	fund ou +	0,79	0,66	0,59	0,01
	até fund,	0,73	0,55	0,47	
AL25	fund ou +	0,79	0,64	0,58	0,01
	até fund,	0,74	0,55	0,48	
AL30	fund ou +	0,81	0,66	0,60	p<0,01
	até fund,	0,72	0,53	0,45	
AL35	fund ou +	0,80	0,63	0,57	0,01
	até fund,	0,75	0,56	0,47	
AL40	fund ou +	0,81	0,66	0,60	0,01
	até fund,	0,75	0,55	0,47	
AL45	fund ou +	0,80	0,65	0,59	p<0,01
	até fund,	0,72	0,54	0,46	
AL50	fund ou +	0,81	0,67	0,61	0,02
	até fund,	0,73	0,55	0,47	

Anexo 14— Probabilidade de sobrevida acumulada do total de pacientes com perdas geradas nos óbitos segundo categoria até o primeiro, terceiro e quinto ano de seguimento, RHC, 1994.

Banco	Categoria	Ano			Teste Log Rank (p)
		1 ano	3 ano	5 ano	
Total	Conv, ou Part	0,72	0,55	0,48	0,01
	SUS	0,69	0,49	0,41	
O5	Conv, ou Part	0,73	0,56	0,50	0,01
	SUS	0,70	0,51	0,44	
O10	Conv, ou Part	0,75	0,58	0,52	0,01
	SUS	0,71	0,52	0,46	
O15	Conv, ou Part	0,76	0,61	0,56	0,01
	SUS	0,73	0,55	0,49	
O20	Conv, ou Part	0,77	0,63	0,57	0,02
	SUS	0,73	0,56	0,50	
O25	Conv, ou Part	0,77	0,62	0,57	0,01
	SUS	0,77	0,59	0,53	
O30	Conv, ou Part	0,77	0,62	0,57	0,00
	SUS	0,77	0,61	0,55	
O35	Conv, ou Part	0,79	0,65	0,60	0,03
	SUS	0,78	0,63	0,56	
O40	Conv, ou Part	0,83	0,70	0,66	0,03
	SUS	0,80	0,65	0,58	
O45	Conv, ou Part	0,84	0,72	0,67	0,00
	SUS	0,81	0,68	0,62	
O50	Conv, ou Part	0,85	0,74	0,71	0,07
	SUS	0,83	0,69	0,65	

Anexo 15– Probabilidade de sobrevida acumulada do total de pacientes com perdas aleatórias segundo escolaridade até o primeiro, terceiro e quinto ano de seguimento, RHC, 1994.

Banco	Categoria	Ano			Teste Log Rank (p)
		1 ano	3 ano	5 ano	
Total	Conv, ou Part	0,78	0,63	0,56	0,00
	SUS	0,73	0,54	0,46	
AL5	Conv, ou Part	0,78	0,63	0,56	0,01
	SUS	0,73	0,54	0,46	
AL10	Conv, ou Part	0,78	0,64	0,57	0,00
	SUS	0,73	0,54	0,46	
AL15	Conv, ou Part	0,78	0,63	0,55	0,00
	SUS	0,73	0,53	0,46	
AL20	Conv, ou Part	0,77	0,63	0,56	0,00
	SUS	0,73	0,54	0,46	
AL25	Conv, ou Part	0,78	0,62	0,56	0,00
	SUS	0,73	0,54	0,47	
AL30	Conv, ou Part	0,77	0,61	0,55	0,00
	SUS	0,72	0,53	0,45	
AL35	Conv, ou Part	0,79	0,62	0,56	0,01
	SUS	0,75	0,55	0,45	
AL40	Conv, ou Part	0,79	0,62	0,56	0,00
	SUS	0,75	0,54	0,47	
AL45	Conv, ou Part	0,79	0,63	0,57	0,00
	SUS	0,71	0,53	0,44	
AL50	Conv, ou Part	0,79	0,63	0,56	0,00
	SUS	0,73	0,55	0,48	

Anexo 16— Probabilidade de sobrevida acumulada do total de pacientes com perdas geradas nos óbitos segundo estadio até o primeiro, terceiro e quinto ano de seguimento, RHC, 1994.

Banco	Estadio	Ano			Teste Log Rank (p)
		3 ano	5 ano		
Total	até 2	0,92	0,82	0,76	p<0,01
	3 ou 4	0,56	0,32	0,25	
O5	até 2	0,84	0,67	0,61	p<0,01
	3 ou 4	0,42	0,22	0,16	
O10	até 2	0,85	0,68	0,62	p<0,01
	3 ou 4	0,45	0,23	0,18	
O15	até 2	0,86	0,71	0,65	p<0,01
	3 ou 4	0,47	0,25	0,20	
O20	até 2	0,86	0,72	0,66	p<0,01
	3 ou 4	0,48	0,28	0,22	
O25	até 2	0,88	0,74	0,69	p<0,01
	3 ou 4	0,49	0,27	0,23	
O30	até 2	0,88	0,75	0,71	p<0,01
	3 ou 4	0,52	0,31	0,24	
O35	até 2	0,89	0,77	0,71	p<0,01
	3 ou 4	0,55	0,35	0,28	
O40	até 2	0,90	0,77	0,72	p<0,01
	3 ou 4	0,59	0,39	0,32	
O45	até 2	0,91	0,80	0,75	p<0,01
	3 ou 4	0,61	0,40	0,34	
O50	até 2	0,91	0,81	0,78	p<0,01
	3 ou 4	0,63	0,43	0,37	

Anexo 17– Probabilidade de sobrevida acumulada do total de pacientes com perdas aleatórias segundo estadio até o primeiro, terceiro e quinto ano de seguimento, RHC, 1994.

Banco	Estadio	Ano			Teste Log Rank (p)
		1 ano	3 ano	5 ano	
Total	até 2	0,94	0,85	0,79	p<0,01
	3 ou 4	0,62	0,38	0,30	
AL5	até 2	0,94	0,85	0,79	p<0,01
	3 ou 4	0,62	0,38	0,30	
AL10	até 2	0,94	0,85	0,79	p<0,01
	3 ou 4	0,61	0,38	0,30	
AL15	até 2	0,94	0,85	0,79	p<0,01
	3 ou 4	0,61	0,38	0,30	
AL20	até 2	0,94	0,86	0,81	p<0,01
	3 ou 4	0,61	0,38	0,30	
AL25	até 2	0,93	0,85	0,80	p<0,01
	3 ou 4	0,62	0,38	0,31	
AL30	até 2	0,93	0,84	0,78	p<0,01
	3 ou 4	0,61	0,37	0,30	
AL35	até 2	0,95	0,87	0,80	p<0,01
	3 ou 4	0,65	0,40	0,32	
AL40	até 2	0,94	0,84	0,78	p<0,01
	3 ou 4	0,63	0,38	0,31	
AL45	até 2	0,93	0,84	0,78	p<0,01
	3 ou 4	0,61	0,37	0,30	
AL50	até 2	0,93	0,87	0,82	p<0,01
	3 ou 4	0,63	0,39	0,31	