Segundo tumor maligno primário no pulmão: regressão linear múltipla em casos de carcinoma brônquico primário

Second primary malign tumor of the lungs: multiple linear regression in primary bronchogenic cancer

Mauro Zamboni¹, Cyro Teixeira da Silva Junior², Edson Toscano³,
Walter Roriz⁴, Paulo de Biasi Cordeiro⁵, Gilberto Perez Cardoso⁶

RESUMO

Introdução: tumores primários múltiplos são comuns após tratamento de uma neoplasia primária. **Objetivo:** verificar a importância relativa de variáveis preditivas na explicação para o tempo de surgimento do segundo tumor maligno primário do pulmão em 16 casos que tiveram como tumor maligno primário o carcinoma brônquico. **Material e métodos:** entre 1986 e 2002 no Instituto Nacional do Câncer, Rio de Janeiro e no Hospital Universitário da Universidade Federal Fluminense, Niterói, avaliamos 16 casos em que o carcinoma brônquico foi o tumor primário inicial por meio de um modelo de regressão linear múltipla, após correlação e transformação logarítmica de dados. Tempo de surgimento do segundo tumor maligno primário no pulmão como variável dependente. Variáveis independentes: idade, sexo, carga tabágica e consumo de álcool. **Resultados:** carcinoma escamoso foi o tipo histológico mais encontrado (81,2%) e o sexo masculino o mais acometido (87,5%). Média de tempo para o surgimento do segundo tumor maligno primário no pulmão foi de 24,4 meses (8–46 meses). Individualmente o tabagismo ($b = -0,078$; $t = -0,818$ ($p=0,430$)) foi o melhor preditor do modelo ($R^2 = 0,0955$; R múltiplo = $0,309$; $F = 0,290$, $p=0,878$), mas sem significância estatística. **Conclusão:** No modelo estudado, as variáveis independentes selecionadas (idade, sexo, tabagismo e consumo de álcool) foram incapazes de prever o tempo de surgimento do segundo tumor maligno primário no pulmão.

Descritores: neoplasias pulmonares, fatores de risco, análise multivariada, regressão linear múltipla.

ABSTRACT

Introduction: the development of multiple malignant tumors is common after treatment of a primary neoplasia. **Objective:** the objective of our paper was to identify prognostic factors for the time for development of a second primary malignant lung cancer between 1986 and 2002 at the Federal Fluminense University in Niterói City and Cancer Institute from Rio de Janeiro, Brazil. **Material and methods:** backward multiple linear regression was performed after correlation and logarithmic transformation of data. Independent variables included age, sex, alcohol consumer and tobacco. Dependent variable was time for development of a second primary malignant lung cancer in 16 cases of bronchogenic carcinoma. **Results:** male patients represented 87.5%. Mean age was 61.0 (44-72) years, with a coefficient of variation of 14.9%. Mean of pack years 71.15 and mean of units of alcohol per week of 27.37. 91.4% of patients smoked and 93.2% were alcoholics. Squamous cell carcinoma was the most common histological type. Coefficient of determination with 0.0955; multiple correlation coefficient of 0.309; coefficients of independent variables: sex ($b = 7.98$, ($p=0.573$)); age ($b = 0.295$, $p=0.537$); tobacco ($b = -0.0781$, $p=0.430$) and alcohol consumer ($b = 0.147$, $p=0.620$). F ratio = 0.2904 ($p=0.878$). **Conclusion:** in multiple linear regression analysis, the independent predictors age, sex, alcohol consumer and tobacco was not considered significant for development of a second primary malignant lung cancer.

Keywords: lung neoplasms, risk factors, multivariate analysis, multiple linear regression.

Introdução

A causa dos cânceres primários múltiplos envolvendo um mesmo órgão parece estar relacionada aos efeitos sinérgicos de vários agentes carcinogênicos atuando no crescimento de um grupo celular, alterando-os e criando o chamado “campo de cancerização”¹.

Constituem fatores de risco para câncer do pulmão o tabagismo ativo ou passivo, dieta inadequada, poluição atmosférica, exposição a substâncias carcinógenas, fatores relacionados com o hospedeiro, além de fatores genéticos².

Este estudo foi realizado com objetivo de verificar no conjunto de variáveis independentes quais as que dentre as selecionadas tem importância relativa para o tempo de surgimento do segundo tumor maligno primário no pulmão, em 16 casos, que tiveram como tumor maligno primário o carcinoma brônquico.

Materiais e métodos

Avaliamos 104 casos de segundo tumor maligno primário no pulmão diagnosticados no período entre 1986 e 2002. Dentre estes, em 16 casos (15,4%) o carcinoma brônquico foi o tumor primário.

Delineamento e locais do estudo: estudo primário, observacional, multicêntrico, retrospectivo, de uma série de casos, aprovado pelo Comitê de Ética em Pesquisa da Faculdade de Medicina da Universidade Federal Fluminense, identificados a partir da análise dos prontuários dos pacientes do Hospital do Câncer do Instituto Nacional de Câncer – Ministério da Saúde (INCA/MS), município do Rio de Janeiro e do Hospital Universitário Antônio Pedro (HUAP), município de Niterói, ambos no Estado do Rio de Janeiro que tiveram o diagnóstico comprovado de segundo tumor maligno primário no pulmão.

Critérios de inclusão: o diagnóstico de segundo tumor maligno no pulmão foi definido após os

pacientes serem submetidos a protocolo de avaliação que incluiu anamnese, com ênfase aos fatores de risco e dados demográficos, exame físico, radiografias convencionais do tórax em incidências pósterio-anterior e de perfil e tomografia computadorizada do tórax. Os procedimentos invasivos para a obtenção de amostras foram realizados de acordo com o local no tórax indicado pelo exame imagético (broncoscopia, punção aspirativa com agulha fina, mediastinoscopia, toracoscopia, toracotomia). As amostras colhidas foram submetidas a exames citológico e histopatológico³. Os laudos foram concluídos com base em critérios universais para o diagnóstico de câncer nas amostras de células e/ou tecido.

Os critérios utilizados para diferenciar entre doença primária ou metastática do pulmão foram: lesão única, maior do que 2 cm de diâmetro, localizada próxima ao hilo pulmonar e com evidência citológica ou histológica de origem brônquica, além da presença de lesão endobrônquica com linfadenomegalia peribrônquica, crescimento lento do tumor, e, obviamente, tipos histológicos diferentes. Os critérios imagéticos que favoreceram doença metastática foram: as lesões periféricas múltiplas, o crescimento tumoral rápido e a concordância com o quadro clínico³.

Critérios de exclusão³: a não comprovação do diagnóstico do segundo tumor maligno primário no pulmão ou doença metastática.

Análise estatística⁴⁻⁸: após a coleta, os dados dos 16 casos com carcinoma brônquico primário foram introduzidos numa planilha eletrônica do *Microsoft Excel*[®], versão 2000. No momento da análise estatística, os dados selecionados para cálculo foram transferidos para planilha do programa *MedCalc*, com objetivo de estudo estatístico descritivo e inferencial.

O estudo inferencial foi realizado por meio de análise multivariada com utilização da técnica de regressão linear múltipla.

1. Pneumologista do S. de Tórax do Hospital do Câncer – INCA/MS. Título de Especialista pela Sociedade Brasileira de Pneumologia e Tisiologia. Mestre em Medicina pela Universidade Federal Fluminense.

2. Professor Adjunto do Departamento de Medicina Clínica da Universidade Federal Fluminense. Disciplina de Pneumologia. Título de Especialista pela Sociedade Brasileira de Pneumologia e Tisiologia.

3. Cirurgião de Tórax do Hospital de Câncer – INCA/MS. Título de Especialista pela Sociedade Brasileira de Cirurgia de Tórax.

4. Cirurgião de Tórax do Hospital de Câncer – INCA/MS. Título de Especialista pela Sociedade Brasileira de Cirurgia de Tórax. Livre Docente da Universidade Federal do Rio de Janeiro.

5. Cirurgião de Tórax do Hospital de Câncer – INCA/MS. Título de Especialista pela Sociedade Brasileira de Cirurgia de Tórax. Professor Titular da Universidade Federal Fluminense. Chefe do Serviço de Tórax do Hospital do Câncer – INCA/MS.

6. Professor Titular do Departamento de Medicina Clínica da Universidade Federal Fluminense.

Trabalho realizado em conjunto pelo Programa de Pós-Graduação *Stricto Sensu* em Ciências Médicas da Universidade Federal Fluminense e pelo S. de Tórax do Hospital de Câncer – INCA/MS, Estado do Rio de Janeiro.

Endereço para correspondência: Mauro Zamboni. Rua Sorocaba, 464/302 Botafogo 22271-110 Rio de Janeiro - RJ. Tele/Fax: (21) 2537-5562. E-mail: zamboni@iis.com.br

Artigo recebido para publicação no dia 11/08/2004 e aceito no dia 17/09/2004, após revisão.

A investigação das possíveis variáveis associadas ao tempo de surgimento do segundo tumor maligno primário no pulmão foi realizada utilizando-se os seguintes parâmetros fundamentais para construção de um modelo paramétrico de regressão linear múltipla: 1) variável dependente medida em escala contínua; 2) variáveis independentes medidas em escala contínua ou variáveis categóricas transformadas em variáveis *dummy*; 3) distribuição normal de todas as variáveis e caso esse critério não seja atingido, ajustar os dados das variáveis com transformação algébrica; 4) relação linear entre a variável dependente e as variáveis independentes; 5) não haver interação entre as variáveis independentes e a variável dependente; 6) não ocorrência de multicolinearidade entre as variáveis independentes; 7) variáveis independentes medidas sem erro⁶.

As exigências mencionadas no parágrafo anterior devem ser satisfeitas para se realizarem inferências válidas sobre os coeficientes de regressão, embora esses parâmetros não sejam necessários para calcular tais coeficientes.

O tempo de surgimento do segundo tumor maligno primário no pulmão foi considerado como variável dependente (y), quantitativa, discreta, em meses.

As variáveis independentes avaliadas foram: sexo (x_1), idade (x_2), tabagismo (x_3) e consumo de álcool (x_4).

O sexo dos pacientes foi considerado como variável independente, categórica, nominal, dicotômica (1 = masculino; 0 = feminino). As idades dos pacientes foram consideradas como variável independente, quantitativa, discreta, em anos. A carga tabágica dos pacientes foi considerada como variável independente, quantitativa, discreta, em maços por ano. O consumo de álcool dos pacientes foi considerado como variável independente, quantitativa, discreta, em unidades de álcool consumidas por semana.

A regressão linear múltipla (RLM) foi construída pelo método de regressão *backward*, isto é, todas as variáveis independentes foram colocadas inicialmente no modelo.

Na RLM foi utilizado o critério, proposto por Pagano & Gauvreau⁶, para inclusão de variáveis independentes na modelagem final: provável existência de associação entre cada variável independente e a variável dependente, após análise univariada do coeficiente de correlação, se necessário ajustado com transformação algébrica dos dados, análise de multicolinearidade e normalidade.

Apesar do tamanho da amostra, o coeficiente de correlação de Pearson (r) foi calculado.

Devido o tamanho da amostra, a transformação

logarítmica dos valores das variáveis independentes foi efetuada antes do teste de normalidade. Segundo o teorema do limite central, quanto maior a amostra, maior a chance de que as distribuições das médias das variáveis envolvidas estejam normalmente distribuídas, apesar de não terem individualmente o formato normal. Logo, aumentando-se o tamanho da amostra, os efeitos da não-normalidade das variáveis são reduzidos, aumentando a robustez da análise, e tornando menos necessária a transformação dessas variáveis.

O teste de Kolmogorov-Smirnov foi utilizado para verificar a normalidade dos dados.

Avaliação de multicolinearidade entre as variáveis independentes foi verificada pela presença de coeficientes de correlação elevados na análise univariada. Outro sinal de multicolinearidade verificado foi a instabilidade dos coeficientes estimados e de seus erros-padrão, os quais, particularmente, tornam-se muito elevados.

Para análise do modelo final de RLM foram considerados os coeficientes de regressão, com seus respectivos erros padrão e valores da estatística t para desempenho individual de cada variável independente no modelo. O valor de p foi considerado significativo com 95 % de probabilidade alfa ($p < 0,05$).

Análise de variância (teste F), análise dos resíduos, coeficiente de correlação múltiplo, coeficiente de determinação, coeficiente de determinação ajustado para avaliação do modelo final.

Resultados

A tabela 1 resume os resultados da estatística descritiva dos 16 casos de carcinoma brônquico.

Tabela 1 – Carcinoma brônquico primário. 16 casos relacionados com tempo de surgimento de segundo tumor maligno primário no pulmão.

Descrição	Valores
Média e limite de idade (anos)	61,0 (44 – 72)
Desvio padrão da idade	9,1
Coefficiente de variação da idade (%)	14,9
Média e limite do tempo para o surgimento do segundo tumor primário no pulmão (meses)	24,43 (8,0 – 46,0)
Desvio padrão do tempo surgimento	12,29
Sexo feminino (%)	12,5
Sexo masculino (%)	87,5
Média carga tabágica (maços/ano)	71,15
Média consumo álcool (unidades/semana)	27,37
Carcinoma escamoso (%)	81,2
Adenocarcinoma (%)	18,8

FONTE: INCA/HUAP.

Tabela 2 – Coeficientes de correlação (*r*) das variáveis independentes pesquisadas antes e após transformação logarítmica em relação ao tempo de surgimento do segundo tumor maligno primário no pulmão.

Variáveis independentes	Valor de r	Valor de p	Valor de r após transformação logarítmica	Valor de p
Sexo	0,077	0,7757	-	-
Idade	0,115	0,671	0,106	0,695
Fumo	- 0,136	0,615	- 0,395	0,129
Álcool	0,079	0,7707	0,0777	0,775

FONTE: INCA/HUAP (1986-2000)

Tabela 3 – Análise de normalidade das variáveis pelo Teste de Kolmogorov-Smirnov.

Variáveis	Teste Kolmogorov-Smirnov	Valor de P *
Tempo surgimento segundo tumor maligno primário no pulmão (meses)	151,06	0,884
Sexo	0,12	<0,001
Idade (anos)	82,93	0,704
Fumo (maços/ano)	1867,62	0,938
Consumo de álcool (unid/semana)	275,71	0,583

FONTE: INCA/HUAP (1986-2000)

* Aceita-se normalidade ou H_0 quando $p > 0,05$.

Tabela 4 – Resultados da análise de regressão linear múltipla aplicada aos dados de 16 pacientes com carcinoma brônquico primário. Tempo de surgimento do segundo tumor maligno primário no pulmão como variável dependente.

Variável Independente	Coefficientes de Regressão (b)	Erro Padrão	Valor de t ₍₁₁₎	Valor de p
Intercepto (a)	2,728	-	-	-
Sexo	7,980	13,761	0,580	0,573
Idade	0,295	0,465	0,636	0,537
Fumo	- 0,078	0,095	- 0,818	0,430
Álcool	0,147	0,289	0,509	0,620

Coefficiente de determinação (R^2) = 0,0955; R^2 ajustado = -0,234; coeficientes de correlação múltiplo = 0,309; Desvio padrão residual = 13,649; $F = 0,290$ ($p = 0,878$). $T_{(11)}$: Valor de t com os 11 graus de liberdade do modelo.

O resultado da análise univariada está descrito nas tabelas 2 e 3 e o modelo final com os resultados da regressão linear múltipla na tabela 4.

Discussão

a) Segundo tumor maligno primário no pulmão, tempo de surgimento e tumores primários

A literatura médica internacional tem destacado a ocorrência das neoplasias primárias múltiplas, em diferentes órgãos, numa frequência que varia de 1,7% a 3,9% da população geral⁹.

Diversos trabalhos identificaram o pulmão como o local mais freqüente para o surgimento de um segundo tumor primário nos pacientes com os tumores primários oriundos da cabeça e pescoço (15% e 61,5%) e de outros locais como o próprio pulmão^{10,11}.

O carcinoma brônquico primário apareceu, no presente trabalho, com uma prevalência de 15,4%, sendo o tipo histológico mais prevalente o escamoso (81,2%), assim como o sexo masculino (87,5%).

Teppo e colaboradores¹² concluíram que os tipos histológicos do carcinoma brônquico com maior risco

relativo para surgir um segundo tumor maligno primário no pulmão foram os carcinomas indiferenciados de pequenas células e os adenocarcinomas.

Do ponto de vista clínico, essas observações apontam para a necessidade de serem adotados protocolos adequados à investigação desses pacientes, que incluiriam um exame detalhado dos pulmões já na primeira consulta após tratamento cirúrgico ou outro, medidas contínuas de prevenção, como a interrupção do álcool e do fumo, e controles médicos periódicos¹³⁻¹⁵. Na periodicidade da avaliação, após diagnóstico e tratamento, recomendamos observar que a média de tempo para surgimento do segundo tumor maligno primário no pulmão foi de 24,4 meses, com valor mínimo de 8 meses e máximo de 46 meses em casos de tumor brônquico primário (tabela 1).

b) Análise multivariada por regressão linear múltipla: parâmetros, medidas de associação e de significância estatística

Os testes estatísticos são escolhidos de acordo com os objetivos da pesquisa. A estatística multivariada possui diferentes métodos. Cada análise tem sua fundamentação teórica e sua aplicabilidade¹⁶.

A análise de regressão pode ser Linear Simples (relaciona duas variáveis através de uma reta), Linear Múltipla (relaciona três ou mais variáveis por meio de uma reta) ou Não-Linear (relaciona duas ou mais variáveis por meio de uma curva matemática que não a reta).

A regressão linear múltipla é uma técnica multivariada cuja finalidade principal é obter uma relação matemática entre uma das variáveis (a variável dependente) e o restante das variáveis que descrevem o sistema (variáveis independentes). Sua principal aplicação, após encontrar a relação matemática, é produzir valores para a variável dependente quando se têm as variáveis independentes. Ou seja, ela pode ser usada na predição de resultados¹⁶.

Outras indicações numeradas por Abbad e Torres¹⁷ são: 1) grau de relacionamento entre as variáveis, indicando se uma correlação é significativamente diferente de zero; 2) importância relativa das variáveis preditoras na explicação da variável dependente; 3) magnitude do aumento da correlação múltipla resultante da adição de uma ou mais variáveis na equação; 4) à maneira pela qual uma variável independente se comporta no contexto de outra(s) variável(is); 5) natureza do relacionamento entre as variáveis independentes e dependentes, indicando se o relacionamento é linear ou não-linear; 6) comparação entre conjuntos diferentes de variáveis independentes na predição da variável dependente; 7) cálculo estimativo dos escores da variável dependente para os membros de uma nova amostra ainda não pesquisada; e 8) identificação de relacionamentos causais entre variáveis quando aplicada como um caso especial de *path analysis* ou equação estrutural.

Para que o uso da equação de regressão linear múltipla seja eficaz na predição da variável dependente em estudo, o pesquisador deve examinar previamente os parâmetros da RLM, bem como identificar as consequências da sua violação. Entre os parâmetros citados por Abbad e Torres¹⁷ após citarem Tabachnick e Fidell¹⁸, estão a multicolinearidade, a singularidade, a homogeneidade nas variâncias, a normalidade e a linearidade.

Todos os parâmetros descritos na literatura para apresentação de uma equação com modelo linear múltiplo foram verificados na realização desta pesquisa.

Quando algum dos requisitos para o emprego da estatística paramétrica não puder ser preenchido pelos dados de uma amostra, o pesquisador pode ainda tentar o recurso da transformação dos dados, antes de optar pela aplicação da estatística não-paramétrica. É um recurso que sempre vale a pena tentar, porque a estatística paramétrica é evidentemente mais poderosa que a não-paramétrica¹⁹.

Uma vez normalizada e homogeneizada a distribuição dos dados amostrais, por intermédio da transformação que se comprovar mais conveniente, o pesquisador estará autorizado a utilizar os testes paramétricos. Contudo, se mesmo tendo tentado todos os recursos disponíveis ainda assim a distribuição continua se demonstrando não-normal, ou não-homogênea, ou até mesmo não-aditiva, não há outra alternativa senão utilizar a estatística não-paramétrica¹⁹.

Análise da tabela 2 mostra que, embora os valores individuais de cada variável independente não tenham sido significativo, verifica-se que somente o valor de *p* ficou reduzido após transformação logarítmica, ou seja, produziu um efeito aditivo, embora não significativo, em relação aos dados não ajustados.

Análise da tabela 3 demonstrou que todas as variáveis independentes possuem distribuição normal pelo teste de Kolmogorov-Smirnov, com exceção do sexo que foi definida como variável categórica, dicotômica, que pelos parâmetros (pressupostos) da RLM entrou no modelo de maneira direta. A falta de normalidade geralmente vem junto com falta de homogeneidade de variâncias. Frequentemente, a mesma transformação estabiliza a variância e aproxima para normalidade. Por este motivo é que primeiro foi usado a transformação logarítmica.

Análise da tabela 4 demonstrou que, individualmente, nenhuma variável foi significativa, estatisticamente, para compor um modelo com regras exclusivas. Observa-se que pouca representatividade é exercida por estas variáveis, afinal os coeficientes de regressão (*b*) pouco se afastam de zero. Porém, individualmente, devido ao valor da estatística *t* e seu correspondente valor de *p*, o tabagismo foi o melhor preditor do modelo. Juntamente com os outros preditores, não foi descartado do modelo porque não foi o objetivo do presente estudo.

A qualidade do modelo de investigação adotado pelos autores, resumida na tabela 4, pode ser avaliada por meio do valor do coeficiente de determinação e da distribuição dos resíduos. O coeficiente de determinação (R^2) calculado de 0,0955 significa que os preditores explicaram 9,55% da variância de *y*. Em outras palavras, é a quantidade da variância da variável dependente que é explicada conjuntamente pelas variáveis independentes e é a estatística mais utilizada para interpretar os resultados da regressão¹⁷.

Pode-se concluir, pela análise dos resíduos, que 90,45% da variância da regressão não depende das variáveis independentes estudadas. Isto é, a maior parte da explicação para o tempo de surgimento do segundo tumor maligno primário no pulmão é devido a outras variáveis preditoras não incluídas no modelo,

o que certamente aumentaria o valor do coeficiente de determinação.

O coeficiente de correlação múltiplo encontrado no valor de 0,309 reflete o que foi concluído no parágrafo anterior, ou seja, o moderado grau de associação entre as variáveis preditoras (idade, sexo, tabagismo e consumo de álcool) e o tempo de surgimento do segundo tumor maligno no pulmão. Ainda da tabela 4, o valor de $F = 0,29$ ($p = 0,878$) indicou que não devíamos rejeitar o modelo linear proposto.

O resultado final de uma RLM é uma equação da reta que representa a melhor predição de uma variável dependente a partir de diversas variáveis independentes. Esta equação representa um modelo aditivo, no qual as variáveis preditoras somam-se na explicação da variável critério. A equação da regressão linear pode ser representada por: " $y = a + bx_i + \hat{I}$ ", onde: " y " é a variável dependente; " a " é a constante ou o intercepto entre a reta e o eixo ortogonal; " b " é o coeficiente padronizado de regressão, ou peso; " x_i " são as variáveis independentes (preditoras) e " $\hat{I}$ " é o erro ou resíduo, que se refere à diferença entre os valores observados e preditos¹⁸.

Para finalizar, chamamos a atenção que estas avaliações quantitativas fornecem uma bússola para orientar as medidas preventivas, no caso em direção às variáveis explicativas que mais influenciam o tempo de surgimento do segundo tumor maligno primário no pulmão, e que podem ser anulados ou alterados os seus efeitos deletérios, juntamente com outras variáveis a serem pesquisadas em nosso banco de dados no acompanhamento da evolução de pacientes após tratamento de um carcinoma brônquico primário.

Concluiu-se que em casos de tumor brônquico primário, o tabagismo foi a variável independente preditiva de risco que apresentou o melhor desempenho para explicar o tempo de surgimento do segundo tumor primário maligno no pulmão no modelo de regressão linear múltipla estudado. Outras variáveis preditivas de risco necessitam serem pesquisadas.

Agradecimentos

Os demais autores agradecem ao Professor Dr. Cyro Teixeira da Silva Junior pela análise estatística do trabalho.

REFERÊNCIAS BIBLIOGRÁFICAS

1. Slaughter DP, Southwick HW, Smejkal W. Field cancerization in oral stratified squamous epithelium. *Cancer* 1953;6:963-8.
2. Silva Junior CT, Cardoso GP. Endocrine expression in bronchogenic carcinoma. *Rev Port Pneumol* 2003;IX(2):109-15.
3. Zamboni MM. Tabagismo e consumo de álcool como fatores de risco para o surgimento do segundo tumor maligno primário no pulmão. [Tese]. Niterói (RJ): Universidade Federal Fluminense; 2004.
4. Pereira MG. Epidemiologia – teoria e prática. 3aed. Rio de Janeiro: Guanabara Koogan; 2000.
5. Neufeld JL. Learning business statistics with Microsoft® Excel 2000. USA: Pearson Education, Inc; 2001.
6. Pagano M, Gauvreau K. Principles of biostatistic. 2nd ed. USA: Thomson Learning; 2000.
7. Anderson DR, Sweeney DJ, Williams TA. Essentials of statistics for business and economics. 2nd ed. USA: Thomson Learning; 2003.
8. Katz MH. Multivariable analysis: a primer for readers of medical research. *Ann Intern Med* 2003;138(8):644-50.
9. Thomas GW. The incidence and significance of multiple primary malignant tumors. *Am J Med Sci* 1964;247:427-30.
10. Gluckman JL, Crissman JD. Survival in 548 patients with multiple neoplasms of the upper aerodigestive tract. *Laryngoscope* 1983;93:71-4.
11. Hsieh CW, Chen YM, Perng RP. Temporal relationship between cancers of the lung and aerodigestive tract. *Jpn J Clin Oncol* 1997;27:63-6.
12. Teppo L, Salminen E, Pukkala E. Risk of a new primary cancer among patients with lung cancer of different histological types. *Eur J Cancer* 2001;37(5):613-9.
13. Soria JC, Izzo J, Mao L, Hong WK, Papadimitrakopoulou VA. Biomarkers of carcinogenesis of the upper and lower airway epithelium and the role of chemoprevention. *Bull Cancer* 2001;88(4):351-61.
14. Algranti E, Menezes AM, Achutti AC. Lung cancer in Brazil. *Semin Oncol* 2001;28(2):143-52.
15. Schirmer CC, Gurski R, Castro, MAA. Neoplasias associadas ao carcinoma epidermóide do esôfago. *Rev Assoc Med Bras* 1997;43(4):335-9.
16. Moita Neto JM. Estatística multivariada. Uma visão didática-metodológica. 2004; Disponível em: URL: http://www.criticanarede.com/cien_estatistica.html
17. Abbad G, Torres CV. Regressão múltipla *stepwise* e hierárquica em psicologia organizacional: aplicações, problemas e soluções. *Estud Psicol (Natal)* 2002;7:19-29.
18. Tabachnick B, Fidell L S. Using multivariate statistics. 3rd ed. New York (USA): Harper Collins; 1996.
19. Campos GM. Estatística prática para docentes e pós-graduandos. Transformação dos dados amostrais. 2000; Disponível em: URL: <http://www.forp.usp.br/restauradora/depart.htm> ■