Caracterização e modelo logístico na descrição da hepatite B ou C entre os doadores do Hemocentro Regional de Maringá

Margareth Cizuka Toyama Udo^{1*}, Rosangela Getirana Santana¹, Kioko Elza Takahashi², Márcia de Lima Agosti², Sandra Regina Corbello Pereira² e Zenilda Soares Beltrami²

¹Departamento de Estatística, Universidade Estadual de Maringá, Av. Colombo, 5790, 87020-900, Maringá-Paraná, Brazil.

²Hemocentro Regional de Maringá. *Author for correspondence.

RESUMO. Este artigo apresenta uma aplicação dos métodos estatísticos-medidas de associação e regressão logística para variáveis categorizadas de um estudo comparativo cujo objetivo é identificar fatores de risco para a condição de hepatite B ou C, buscando sanar as deficiências e as distorções no processo de triagem clínica dos candidatos a doadores de sangue do Hemocentro Regional de Maringá, no período de 1993 a 1996. Pela regressão logística, os fatores de risco detectados foram possuir mais de três parceiros sexuais e ter tido parceiros desconhecidos, homossexual ou prostituta. Apesar de o modelo ser significativo, é restrito no que se refere à previsão, uma vez que, pelo critério de classificação, o percentual de acertos entre os doadores com pelo menos uma sorologia positiva de hepatite é baixa. Observa-se que a entrevista médica na triagem clínica trouxe uma contribuição significativa na identificação de inaptidão.

Palavras-chave: medidas de associação, regressão logística, sorologia positiva de hepatite.

ABSTRACT. Characterization and logistic models in hepatitis B or C among the donors of Hemocentro Regional de Maringá. This article presents an application for the statistical methods for classified variables, such as association measures and logistic regression, of a comparative study whose objective is to identify risk factors for the hepatitis B or C condition, correcting deficiencies and distortions in the process of clinical screening of candidates to blood donors at *Hemocentro Regional de Maringá*, in the period from 1993 to 1996. For the logistic regression the detected risk factors were: to have more than three sexual partners and have had unknown partners, homosexual or prostitute. In spite of the model being significant it restricted as for prognosis, once for the classification rule the percentile of successes among the donors with at least one positive serology of hepatitis is low. It was observed that the medical interview in the clinical screening provided a significant contribution to the inaptitude identification.

Key words: association measures, logistic regression, hepatitis positive serology.

O objetivo deste trabalho é caracterizar os candidatos a doadores de sangue, no período de 1993 a 1996, e com isto solucionar um dos problemas encontrados pela equipe do Hemocentro Regional de Maringá, que constatou um alto índice de candidatos a doadores aptos pela triagem clínica e inaptos pela sorológica no ano de 1994; além disso, conhecer o perfil dos candidatos a doadores facilitaria em campanhas educativas e de recrutamento.

Para os candidatos a doadores cadastrados nos anos de 1995 e 1996 no Hemocentro Regional de Maringá, um estudo comparativo foi realizado com o objetivo de detectar fatores de risco para a variável

resposta binária condição de hepatite B ou C. Esse estudo comparativo foi desenvolvido, conforme descrito por Soares e Colosimo (1995), entre os doadores com, pelo menos, uma sorologia positiva de hepatite B ou C e aqueles com todas as negativas.

A análise de estudos comparativos se faz primeiramente identificando associações para cada fator em estudo, conforme Agresti (1990), e posteriormente analisando os fatores conjuntamente, que é objeto da regressão logística, dado que a variável resposta é binária.

O procedimento de regressão logística foi primeiramente adotado por Truett *et al.* (1967), na análise de dados do experimento de Framingham,

que tinha como objetivo fornecer uma análise multivariada de fatores de risco em pacientes com problemas cardíacos. Esse modelo é considerado padrão na análise de variável binária em diversos experimentos principalmente nas pesquisas clínicas, e assegura que qualquer estimativa de probabilidade obtida será sempre algum número entre 0 e 1. Isto não é sempre verdadeiro para outros possíveis modelos, e é por isso que o modelo logístico é o mais adequado quando o objetivo é estimar a probabilidade de um evento. Neste estudo, tal probabilidade fornece o risco de um candidato a doador de sangue apresentar, pelo menos, uma sorologia positiva de hepatite B ou C.

Segundo Hosmer e Lemeshow (1989), a qualidade dos artigos publicados que utilizam o método de regressão logística não era a mesma da regressão linear, contudo, a inclusão do procedimento de regressão logística em *software* estatísticos, tais como GLIM, Statistica e SAS, permitiu que o método se tornasse mais conhecido.

Assim, este estudo foi desenvolvido em duas etapas, sendo que, na primeira, identificaram-se as principais causas de inaptidão e descreveu-se a condição de aptidão através do perfil dos doadores. A análise descritiva dos dados coletados revelou que a principal causa de inaptidão era, pelo menos, uma sorologia positiva de hepatite, que vinha aumentando ano a ano. Diante disso, passou-se para a segunda etapa, que consistiu em identificar os fatores de risco pelos testes de homogeneidade, independência e medidas de risco, conforme descrito em Goodman (1972), e obter um modelo logístico que descreve a variável resposta, condição de hepatite B ou C.

Metodologia

A população consistiu de todos os candidatos que se apresentaram para a doação de sangue nos anos de 1993 a 1996. Esses candidatos são, inicialmente, cadastrados no Livro de Registro de Doadores, e as informações prestadas à recepcionista são anotadas na Ficha Cadastral. Em seguida, eles passam por dois tipos de triagem: a clínica, que identifica se o candidato tem algum problema de saúde, como anemia, febre, baixo peso, hipertensão, etc., e a sorológica, que avalia se o sangue coletado tem algum tipo de contaminação, como hepatite, malária, HIV, sífilis e HTLV. A partir de 1995, além dessas duas triagens, o candidato a doador passou a ser entrevistado com questões elaboradas pelos profissionais do Hemocentro, consideradas relevantes, segundo Di Colli (1995), na avaliação da hepatite B ou C.

Um arquivo de dados, contendo as informações das fichas cadastrais e da entrevista médica, foi criado no *software Epidemiology Program Office* (EPI6) para análise estatística. O procedimento consistiu em descrever as variáveis relacionadas nas fichas cadastrais, obtendo o perfil dos candidatos a doadores, verificar as associações entre variáveis, e identificar variáveis que explicam a condição de, pelo menos, uma sorologia positiva de hepatite pelo modelo de regressão logística, uma vez que essa variável resposta é do tipo binária. Para a obtenção das estatísticas, dos testes e dos modelos, utilizamos os *softwares* Statistica (1994) e EPI6.

Como no delineamento experimental, somente o total de candidatos a doadores é fixado, a *priori*, a distribuição da variável resposta é a multinomial, assim, o enfoque não condicional para os testes estatísticos foi adotado, com nível de significância de 5%.

Para a categorização das variáveis quantitativas, adotou-se o método de conglomerado K-médias, e para a obtenção das variáveis indicadoras, identificadas, através do risco relativo, a categoria 1, que representa o exposto, e a 0, que representa o não-exposto.

Resultados e discussão

O perfil mostra que 74% dos candidatos são homens, com idade média de 32 anos e com desvio-padrão de 10 anos, sendo a maioria da raça branca (85%). Dos candidatos, 33,5% são considerados inaptos, tendo como principal causa de inaptidão temporária a anemia e, definitiva, a hepatite B ou C, que nestes quatro anos tem aumentado significativamente.

Pela análise univariada, as variáveis *sexo* e *idade* estão associadas significativamente à condição de aptidão.

Com relação à variável resposta, condição de hepatite B ou C, detectou-se uma associação significativa para as variáveis naturalidade, estado civil, idade, número de parceiros sexuais e parceiros sexuais desconhecido, homossexual ou prostituta (na Tabela 1, é fornecido o risco para essas variáveis). Portanto, os candidatos a doadores que são da região Norte ou Nordeste e/ou com mais de 35 anos e/ou casados ou amasiados e/ou os que têm mais de três parceiros sexuais e/ou aqueles que têm parceiros sexuais desconhecidos, homossexual ou prostituta, têm maior risco de apresentar, pelo menos, uma sorologia positiva de hepatite. As outras variáveis em estudo não foram significativas.

O modelo de regressão logística mostra que as variáveis número de parceiros sexuais e parceiros sexuais desconhecidos, homossexual ou prostituta

contribuíram significativamente na explicação de, pelo menos, uma sorologia positiva de hepatite, conforme mostrado na Tabela 2.

Tabela 1. Análise univariada para a hepatite B ou C por variáveis clínicas Hemocentro Regional de Maringá - 1996

Variável	p	Risco Relativo Intervalo de 95% de Confiança
Naturalidade:		
0 - sul, sudeste, centro - oeste	0,00146	1,16 (1,02 ; 1,33)
1 - norte, nordeste		
Estado Civil:		
0 - solteiro, divorciado e viúvo	0,00200	1,07 (1,03 ; 1,11)
1 - Casado, amasiado		
Idade		
0 - até 35 anos	0,00000	1,09 (1,05 ; 1,13)
1 - mais 35 anos		
Número de parceiros sexual no último ano:		
0 - menos de 3	0,00070	1,13 (1,06 ; 1,20)
1 - mais de 3		
Relacionamento com parceiro desconhecido e/ou homossexual e/ou prostituta, no último ano:		
0 - Não	0,00383	1,08 (1,02 ; 1,15)
1 - Sim		

Tabela 2. Significância do modelo, estimativa dos parâmetros, desvio-padrão e significância dos parâmetros do modelo de regressão logística

Variável Dependente: Sorologias de Hepatite			
Perda Final :106,015 $\chi^2 = 11,448$ P=0,04321			
Variável	Parâmetros Estimados	Desvio-Padrão	Valor de P*
Naturalidade	0,334402	0,272191	0,219346
Idade	-0,04708	0,409122	0,908394
Número de parceiros sexual	0,473962	0,123433	0,000126
Relacionamento com parceiro desconhecido e/ou homossexual e/ou prostituta, no último ano	1,226090	0,623487	0,049345

*Nível de significância do parâmetro do modelo pelo teste t - Student

Pela Tabela 3, o critério de classificação que avalia o modelo de regressão logística mostra que não é boa para previsão, uma vez que só consegue classificar os não-portadores de hepatite. Isto se deve, principalmente, ao fato de a hepatite ser uma doença silenciosa, já que não apresenta sintomas na maioria dos casos e, assim, não é percebida pelo portador, sendo o diagnóstico definido somente na triagem sorológica.

Tabela 3. Classificação dos casos com relação às sorologias de hepatite

Observado	Estimado		Percentual Correto
	Negativa	Positiva	
Negativa	2110	234	90
Positiva	202	109	35

Pelas Tabelas 4 e 5, a entrevista médica apresentou bons resultados, visto que, a partir dela, a prevalência na identificação da inaptidão e de pelo menos uma sorologia positiva de hepatite aumentou

em 0,32 e 0,27, respectivamente. Desses valores, conclui-se o que algumas deficiências na triagem clínica foram sanadas.

Tabela 4. Distribuição por entrevista médica da condição de aptidão

Entrevista	Condição de Aptidão		Total
	Inapto	Apto	
Sem	333	1510	1843
Com	558	3526	4084
Total	891	5036	5927
$\chi^2 = 19,29$ P = 0,0000112			
Prevalência = 1,32 (1,17 ; 1,50)			

Tabela 5. Distribuição por entrevista médica e pelo menos uma sorologia positiva de hepatite

Entrevista	Condição de Aptidão		Total
	Inapto	Apto	
Sem	301	1453	1754
Com	538	3430	3968
Total	839	4883	5722
$\chi^2 = 12,61$ P = 0,00038			
Prevalência = 1,27 (1,11 ; 1,44)			

As variáveis naturalidade, estado civil, idade, três ou mais parceiros no último ano, parceiro desconhecido no último ano, homossexual ou prostituta são fatores de risco para, pelo menos, uma sorologia positiva de hepatite B ou C.

O modelo de regressão logística, apesar de significativo, é restrito no que se refere à previsão de um candidato ser mais susceptível a, pelo menos, uma sorologia positiva de hepatite, uma vez que este classifica corretamente 90% dos candidatos com sorologias negativas de hepatite e 35% dos portadores da doença. Isto se deve principalmente ao fato de a hepatite ser uma doença silenciosa, já que não apresenta sintomas na maioria dos casos e, assim, não é percebida pelo portador, sendo o diagnóstico definido somente na triagem sorológica.

Por último, uma avaliação no processo de triagem clínica mostra que a entrevista com o candidato a doador contribui significativamente na identificação da inaptidão e de, pelo menos, uma sorologia positiva de hepatite.

Referências bibliográficas

- Agresti, A. *Categorical data analysis*. New York: John Wiley and Sons, 1990.
- Di Colli, L. *Prevalência dos marcadores HbSAg e anti-Hbc da hepatite B em doadores de sangue do núcleo de hemoterapia de Apucarana*. Maringá, 1995. Monografia.
- Goodman, L. A. *Some multiplicative models for the analysis of cross-classified data*. Berkley: University of California Press. 1972.

Hosmer, D.W.; Lemeshow, S. *Applied logistic regression*. New York: John Wiley and Sons, 1989.

Soares, J.F., Colosimo, E.A. *Métodos estatísticos na pesquisa clínica*. SIMPÓSIO DE ESTATÍSTICA APLICADA À EXPERIMENTAÇÃO AGRONÔMICA. 1995. Ribeirão Preto. *Anais...* Ribeirão Preto: RBRAS, 1995. p.105.

Statistica: manual. Tulsa: StatSoft, 1994.

Um sistema de processamento de texto, banco de dados e estatística para epidemiologia em micro computadores: manual. Georgia: Epidemiology Program Office, 1990.

Received on October 22, 1999.

Accepted on November 29, 1999.