

**ESTUDOS TRANSVERSAIS E LONGITUDINAIS COM DESFECHOS BINÁRIOS:
QUAL A MELHOR MEDIDA DE EFEITO A SER UTILIZADA?**

*CROSS-SECTIONAL AND LONGITUDINAL STUDIES WITH BINARY OUTCOMES:
WHAT IS THE BEST MEASURE OF EFFECT TO BE USED?*

Vânia Naomi Hirakata

RESUMO

Desde a década de 90, o uso da razão de chances em estudos transversais e longitudinais tem sido discutido na literatura epidemiológica. Este artigo tem como objetivo descrever os principais argumentos trazidos pelos autores a favor e contra o uso da razão de chances nestes tipos de estudos e apresentar as principais alternativas propostas para que seja possível estimar razões de prevalências ou de incidências cumulativas ajustando para demais variáveis.

Unitermos: *Transversal; longitudinal; razão de chances; razão de prevalência; incidência cumulativa*

ABSTRACT

Since the 90s, using the odds ratio in cross-sectional and longitudinal studies have been discussed in the epidemiological literature. This article aims to describe the main arguments brought by the authors for and against the use of odds ratio in these types of studies and present the main alternatives proposed in order to estimate the prevalence ratios or cumulative effects, adjusting for other variables.

Keywords: *Cross-sectional; longitudinal; prevalence ratios; cumulative effects*

Rev HCPA 2009;29(2):174-176

Os estudos epidemiológicos encontrados na literatura usam, com grande frequência, delineamento transversal, por se tratar de uma alternativa rápida, simples e de baixo custo. A partir de um levantamento realizado em um ano de publicação de duas revistas da literatura epidemiológica (*International Journal of Epidemiology* e *Revista de Saúde Pública*, ano de 1998), verificou-se que, entre os 221 artigos originais, 110 (50%) eram baseados em estudos com delineamento transversal e 45 (20%) possuíam delineamento longitudinal (1). Muitas vezes os desfechos são variáveis binárias do tipo óbito/não óbito ou doença/não doença (2,3) e a razão de chances é relatada com muito mais frequência do que a razão de prevalências (2,4).

O alto número de publicações baseadas em estudos transversais e longitudinais que utilizam a regressão logística como ferramenta de análise chama a atenção. A medida de efeito que tem sido estimada nestes casos é a razão de chances, enquanto que em estudos transversais, a medida de efeito de interesse é a razão de prevalências e, em estudos longitudinais (inclusive ensaios clínicos randomizados), a razão de incidências cumulativas (ou risco relativo). Embora a razão de chances seja estimável em todos os delineamentos, o risco relativo é preferível, por diversos motivos (2,5-8).

A maior crítica em relação ao uso da razão de chances é que ela é uma medida de efeito de difícil interpretação (2,7,8). Além disso, muitas vezes é interpretada de forma errada, como um risco relativo (3). Na verdade, o *odds* é a probabilidade de um evento ocorrer, dividido pela

probabilidade dele não ocorrer. A razão de chances, por sua vez, é a razão entre o *odds* dos pacientes expostos (por exemplo, a uma droga) e o *odds* dos pacientes não expostos.

Quando se trabalha com um desfecho raro (prevalência de até 10%), a razão de chances é numericamente similar ao risco relativo e até pode ser interpretada como tal. Porém, quando a prevalência do desfecho é alta ou a magnitude do efeito elevada, a razão de chances superestima o risco relativo, distorcendo, portanto, a magnitude do efeito que se deseja estimar (4,9,10). Portanto, no contexto da saúde pública e da epidemiologia clínica, não é adequado interpretar indiscriminadamente uma razão de chances da mesma forma que um risco relativo (9), embora muitas vezes seja interpretada de forma errada (3,11).

Neste contexto, a regressão logística acaba sendo a técnica mais utilizada quando se deseja analisar múltiplos fatores de risco simultaneamente ou controlar possíveis variáveis de confundimento. Isto ocorre independentemente do tipo de delineamento, por haver uma grande disponibilidade de programas computacionais que a incorporam (2,5,7). Porém, cabe lembrar que a regressão logística utiliza como medida de efeito a razão de chances e não a razão de prevalências ou de incidências, medidas naturais de estudos transversais e longitudinais.

Além disso, sabe-se que diferentes padrões de confundimento podem ocorrer quando calculados o risco relativo e a razão de chances

para os mesmos dados (7,12), fazendo com que o pesquisador tenha conclusões diferentes a respeito de variáveis de confundimento dependendo da medida de efeito que se deseja estimar.

Por estas razões, a regressão logística é considerada uma técnica de análise que deveria ser usada com cautela em estudos transversais e longitudinais.

Alternativas para a análise de dados que estimam o risco relativo em estudos com delineamentos transversal ou longitudinal têm sido propostas na literatura. A regressão de Cox poderia ser utilizada para estimar a razão de prevalências em estudos transversais, pois quando se presume um período de risco constante (coorte fechada de pessoas em risco), a razão de taxas de falhas (*hazard ratio rate*) con-

dicional estimada pela regressão de Cox é igual à razão de incidência cumulativa (2,13). Outra alternativa, no contexto de modelos lineares generalizados baseados em distribuições binomiais, seria a utilização do logaritmo da proporção como uma função de ligação no modelo de regressão, em vez do logito, utilizado na regressão logística. Neste caso, os coeficientes da regressão estimam o risco relativo, controlando para possíveis variáveis de confundimento e considerando interações (14).

Diversas alternativas para a análise de dados que estimam o risco relativo em estudos com delineamentos transversal ou longitudinal têm sido propostas na literatura. A abordagem mais simples é o uso de fórmulas para converter a razão de chances (RO) em risco relativo (RR). Uma das fórmulas de conversão proposta, baseada nas marginais, é dada por:

$$RR = \frac{P_E + P_D - 1 + RQ(P_E - P_D) + \sqrt{[1 - P_E - P_D - RQ(P_E - P_D)]^2 - 4P_E RQ(P_E - 1)}}{2P_E} \quad (1)$$

Onde P_E é a prevalência de exposição e P_D a prevalência de doença. Esta conversão é uma inversão da equação apresentada por Zocchetti(5) que, originalmente, era assim composta:

$$RO = RR \frac{(1 - P_E + RR \cdot P_E - P_D)}{(1 - P_E + RR \cdot P_E - RP \cdot P_D)} \quad (2)$$

Nota-se que, nesta fórmula são necessárias somente as medidas dos valores marginais, ou seja, da prevalência do desfecho e da prevalência das variáveis de exposição, além da razão de chances. São valores diretos e podem ser obtidas facilmente num estudo transversal. Porém, se o delineamento do estudo não permitir as estimativas das prevalências do desfecho e das variáveis de exposição, este cálculo pode não ter validade.

Outra fórmula de conversão é a proposta por Zhang ET AL (15) é:

$$RR = \frac{OR}{(1 - P_0) + (P_0 \times OR)} \quad (3)$$

Onde P_0 é a prevalência/incidência do desfecho entre os não expostos. Esta fórmula pode ser utilizada para converter os valores de razão de chances estimadas pela regressão logística e, assim, obter uma medida de efeito que represente melhor o verdadeiro risco relativo.

A utilização de diversos modelos cujos coeficientes estimem diretamente o risco relativo também têm sido apresentadas na literatura. O modelo log-binomial seria o que estima o risco relativo de forma mais direta, por ser um modelo linear generalizado na qual a função de ligação é o logaritmo da proporção em estudo e a distribuição do erro é binomial. No entanto, diversos autores têm apontado para a falta de convergência, situação na qual o modelo de máxima verossimilhança não consegue estimar um resultado, em diversas situações com este modelo (1,16-18), enquanto outros têm tentado resolver este problema e apontam que este seria o mo-

delo ideal para estimar razões de prevalência (19).

Outro modelo que tem sido testado como alternativa à regressão logística é a regressão de Poisson. O problema desta modelagem era a estimativa de erro, que segue uma distribuição de Poisson, gerando erros-padrões maiores dos que os gerados no modelo log-binomial. No entanto, este problema foi resolvido com o uso de estimativas robustas nas variâncias (1,16,17,20).

As fórmulas de conversão eram as únicas alternativas quando os pacotes estatísticos disponibilizavam apenas a regressão logística.

Atualmente, com a incorporação dos modelos lineares generalizados (*generalized linear models*) é possível realizar as análises multivariáveis por regressão de Poisson robusta ou por regressão log-binomial em pacotes como o SPSS, SAS, R e Stata (1,16,17).

Portanto, o uso da razão de chances, como medida de efeito, deve ser evitada em estudos transversais e longitudinais, principalmente quando o desfecho for frequente.

REFERÊNCIAS

- Barros AJ, Hirakata VN. Alternatives for logistic regression in cross-sectional studies: an empirical comparison of models that directly estimate the prevalence ratio. *BMC Med Res Methodol*. 2003 Oct 20;3:21.
- Lee J. Odds Ratio or Relative Risk for Cross-Sectional Data. *Int J Epidemiol*. 1994;23:201-3.
- Thompson ML, Myers JE, Kriebel D. Prevalence odds ratio or prevalence ratio in the analysis of cross sectional data: what is to be done? *Occup Environ Med*. 1998 Apr;55(4):272-7.
- Nurminen M. To use or not to use the odds ratio in epidemiologic analyses. *Eur J Epidemiol*. 1995;11:365-71.
- Zocchetti C, Consonni D, Bertazzi P. Relationship between Prevalence Rate Ratios and Odds Ratios in Cross-Sectional Studies. *Int J Epidemiol*. 1997;26:220-3.
- Martuzzi M, Elliott P. Estimating the incidence rate ratio in cross-sectional studies using a simple alternative to logistic regression. *Ann Epidemiol*. 1998;8:52-5.
- Axelsson O, Fredriksson M, Ekberg K. Use of the Prevalence Ratio v the Prevalence Odds Ratio as a Measure of Risk in Cross-Sectional Studies. *Occup Environ Med*. 1994;51:574.
- Sackett D, Deeks J, Altman D. Down with Odds Ratios! *Evidence-Based Medicine*. 1996:164-6.
- Greenland S. Interpretation and Choice of Effect Measures in Epidemiologic Analyses. *Am J Epidemiol*. 1987;1987:761-8.
- Altman DG, Deeks JJ, Sackett DL. Odds ratios should be avoided when events are common. *BMJ (Clinical research ed)*. 1998 Nov 7;317(7168):1318.
- Savitz D. Measurements, Estimates, and Inferences in Reporting Epidemiologic Study Results. *Am J Epidemiol*. 1992;135(3):223-4.
- Miettinen O, Cook E. Confounding: Essence and Detection. *Am J Epidemiol*. 1981;114(4):593-603.
- Breslow N. Covariance Analysis of Censored Survival Data. *Biometrics*. 1974;30:89-99.
- Wacholder S. Binomial Regression in GLIM: Estimating Risk Ratios and Risk Differences. *Am J Epidemiol*. 1986;123:174-84.
- Zhang J, Yu KF. What's the Relative Risk? *JAMA*. 1998;280(19):1690-1.
- Coutinho LM, Sczufca M, Menezes PR. Methods for estimating prevalence ratios in cross-sectional studies. *Rev Saude Publica*. 2008 Dec;42(6):992-8.
- Spiegelman D, Hertzmark E. Easy SAS calculations for risk or prevalence ratios and differences. *Am J Epidemiol*. 2005 Aug 1;162(3):199-200.
- Skov T, Deddens J, Petersen MR, Endahl L. Prevalence proportion ratios: estimation and hypothesis testing. *Int J Epidemiol*. 1998 Feb;27(1):91-5.
- Petersen MR, Deddens JA. A comparison of two methods for estimating prevalence ratios. *BMC Med Res Methodol*. 2008;8:9.
- Zou G. A modified poisson regression approach to prospective studies with binary data. *Am J Epidemiol*. 2004 Apr 1;159(7):702-6.

Recebido: 11/06/2009

Aceito: 27/07/2009