Letters to the Editor

RE: “EASY SAS CALCULATIONS FOR RISK OR PREVALENCE RATIOS AND DIFFERENCES”

We applaud Drs. Spiegelman and Hertzmark’s idea of using SAS procedure PROC GENMOD to estimate the risk ratio or difference (1). However, we have reservations about 1) the claim that there is no good justification for fitting the logistic regression and estimating the odds ratio when the odds ratio is not a good approximation of the risk ratio, and 2) using Poisson regression (PROC GENMOD) to estimate the risk ratio when the log-binomial model fails to converge.

In our opinion, the choice of models (e.g., logistic vs. log-binomial regression) is dictated by the data; that is, only that model supported by the data should be used. In this regard, the popularity of the logistic regression model lies in its ability to fit a wide range of data well, rather than the fact that the odds ratio sometimes is an approximation of the risk ratio. In addition, the odds ratio given by the logistic regression model is a good summary of the association in its own right. Furthermore, the failure of convergence in the log-binomial regression is not only a numerical problem but also an indication that the data do not support the model.

We use a simple example from Hosmer and Lemeshow (2) to illustrate our point. The data set consists of 100 participants aged 20–69 years and their coronary heart disease status (presence vs. absence). The goal is to study the relation between the prevalence of coronary heart disease and age. Since the average prevalence in the cohort was 43 percent, the risk ratio and odds ratio were not similar. The simple log-binomial model fails to converge. If we use Poisson regression to estimate the risk ratio (prevalence ratio) in the log-binomial model, the estimated risk ratio for a 10-year increase is 1.72. As shown in figure 1, we compared the resulting model-based prevalence of coronary heart disease with its nonparametric counterpart, estimated from a model assuming only that the prevalence varies smoothly with age. Since nonparametric fitting requires no specific parametric model, the estimated prevalence may serve as the “true” prevalence. The prevalence estimated by the log-binomial model approximates the “true” prevalence well for age ≤60 years but poorly for age >60 years (figure 1). For eight participants older than age 60 years, the average prevalence estimated by using the log-binomial model is 102 percent (meaningless); the observed prevalence is 87.5 percent. In this case, the log-binomial regression model fails to converge because of poor fitting of the data. By forcing...
a constant risk ratio over the entire age range, the iterative algorithm fitting the log-binomial model yields probabilities greater than 1 and stops. In this example, the prevalence estimated by the logistic regression model approximates the “true” prevalence remarkably well (figure 1). In contrast to that based on the log-binomial model, the average estimated prevalence based on logistic regression is 85.0 percent for participants older than age 60 years. Lastly, we also fit a linear regression model, which gave as poor a fitting of prevalence for younger participants as log-binomial regression did for older participants.

We hope that this example demonstrates that there are situations in which the log-binomial model (assuming a constant risk ratio) does not fit the data well, whereas the logistic regression model (assuming a constant odds ratio) does. The first sign of poor fitting of the log-binomial model is its failure to converge. When using Poisson regression to circumvent the convergence problem in fitting the log-binomial model, one ignores the constrictions of model parameters and likely produces meaningless probability. When this situation occurs (not rare in our experience), no matter how appealing the interpretation of the risk ratio, one should conduct logistic or other regression analyses equipped with appropriate model-checking procedures.

**ACKNOWLEDGMENTS**

Conflict of interest: none declared.

**REFERENCES**


L. Tian and K. Liu (e-mail: lutian@northwestern.edu)
Department of Preventive Medicine, Feinberg School of Medicine, Northwestern University, Chicago, IL 60611

DOI: 10.1093/aje/kwj163; Advance Access publication May 17, 2006