Risk factors for respiratory infections amongst preschool children in Indonesia

Georgina Davies
University of Queensland

My project involved analysing a dataset collected in Indonesia in 2003 on the incidence of childhood diseases. The dataset comprised observations on 826 children and the participants were randomly assigned to the zinc or placebo group and all children received vitamin A supplementation halfway through the four month observation period. The diseases of interest for this project were diarrhoea, skin disease and upper respiratory tract infection (URTI).

This dataset was of particular statistical interest because it contained numerous zero responses, because for each disease, there were many children who did not contract it. A zero-altered dataset is one that has more (zero-inflated) or less (zero-deflated) zero responses than would be expected under a typical probability distribution. Zero-inflated count data is common in epidemiological studies and we were interested in exploring the use of the zero-altered Poisson random effect model (Min and Agresti, 2005). It is a special case of a hurdle model. The probability of a non-zero response (contracting the disease) is p, and the non-zero responses follow a truncated Poisson distribution. Mathematically the model is:

$$P(Y_i = 0) = 1 - p$$

$$P(Y_i = y) = p\frac{e^{-\mu}\mu^{y}}{y!(1-e^{-\mu})}$$

The parameters, p and μ are expressed as functions of the covariates (β), and γ is a constant that determines if the dataset is significantly zero altered and r_i is the random effect associated with the ith person:

$$\log\mu = x\beta + r_i$$

$$\log(-\log(1-p)) = \gamma + x\beta + r_i$$

The SAS program was used to fit the model and it was determined that the dataset was significantly zero-inflated for all three diseases; hence the use of zero-altered model is justified. The key clinical results were that in terms of skin disease and URTI the zinc and placebo groups responded differently to vitamin A

supplementation, and for URTI the combination of both zinc and vitamin A was the optimum treatment.

My project has given me experience in consulting with non-statisticians on the analysis of a dataset. Further, it has given me more exposure to the collaborative process often involved in scientific work and journal article writing and expanded my knowledge of statistical models and software packages. I would like to thank my supervisors; Dr Mark Griffin, Dr Michael Bulmer and Assoc Prof Geoff Marks, the University of Queensland, AMSI and CSIRO for their support.

References

Min Y and Agresti A. Random effect models for repeated measures of zero-inflated count data. *Statistical Modelling* 2005;**5**:1-19.

Georgina received an AMSI Vacation Scholarship in December 2010 See: www.amsi.org.au/vs10.php