

Paper PO-24

Breastfeeding in Developing Countries: A Case Study of Nepalese Children

Paerwen Paerhate, George Mason University, Fairfax, VA

ABSTRACT

Breastfeeding is beneficial for the health of the child and the breast feeding mother. It is the recommended practice by the American Academy of Pediatrics. Many studies have been conducted on the effect of breastfeeding on children's health in the developed countries, whereas little is known about breastfeeding in the underdeveloped countries. This study characterized the relationship between breastfeeding and children's age and gender in Nepalese children. Using secondary data that were collected on 200 children at 5 time points, spaced approximately 4 months apart (1000 total observations), a logistic regression model was developed to regress the breastfeeding status on children's age and gender. This study utilized the Generalized Estimating Equations (GEE) from GENMOD procedure in SAS® 9.2 to examine three situations of the distribution of variance of the model error term: compound symmetry (QIC = 539.6898), autoregressive (1) (QIC = 525.8675), and independent (QIC = 523.3782). Preliminary results suggest that the independent covariance structure fits the data best. The study found no significant gender effect on breastfeeding status. The probability of breastfeeding was negatively associated with increasing age of the child but the magnitude of this effect increasingly diminished as age increased.

Introduction

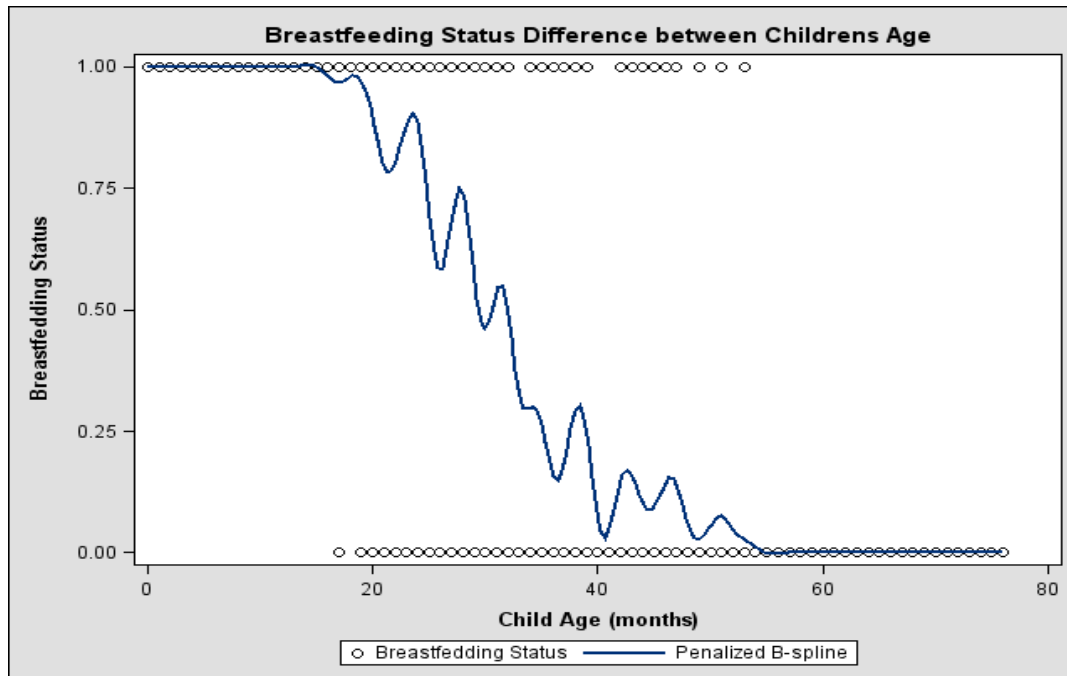
Breastfeeding has been widely considered as important nutrition to keep children healthy. As natural immunization, breastfeeding could prevent children from certain high risk diseases. However, in many underdeveloped countries, breastfeeding is not universally practiced. In Nepal, only 53% of children under six months are exclusively breastfed, 49% of children under five suffer from chronic malnutrition. This paper examined how breastfeeding is related to children's gender and age. A quadratic logistic regression was performed by using GENMOD procedure; three types of covariance structure of error term were tested for goodness of fit. Results show that for those who practiced breastfeeding, women usually stop breastfeeding when child is one year old. The probability of breastfeeding decreases as children grow. This study indicates that GENMOD is a strong procedure which could identify a better covariance structure of generalized models.

Data Summary

This study is based on a secondary dataset which contains anthropologic measurements. The original data contains 9552 complete records on 2215 children, and first 1000 records were extracted for this study. There are 200 children at 5 points, spaced approximately 4 months. After deleting missing data, 947 observations are used for data analysis.

Graphic Analysis

A scatter plot with penalize B-spline curve is performed to get an intuitive relation between children's age and the main outcome-breastfeeding status. SGPLOT procedure is used to generate the following graph:



```
Proc Sgplot Data=nepalbin;
Title 'Breastfeeding Status Difference between Childrens Age';
Scatter x=age y=bfstatus;
Pbspline x=age y=bfstatus;
Label bfstatus='Breastfeeding Status' age='Child Age (months)' ;
run;
```

This graph shows that mothers usually practice breastfeeding while children are less than one year old, as children's age increase, mothers intended to stop breastfeeding.

Logistic Regression and Results

The main outcome breastfeeding is an indicator variable with value (0, 1), for each observation y_{ij} , it is assumed to have Bernoulli distribution with probability p_i . Here i is the number of children range from 1 to 200, j is the number of visit range from 1 to 5. The following is the logistic regression model:

$$\text{Logit}(p_i) = \beta_0 + \beta_1 * \text{male}_i + \beta_2 * \text{age}_{ij} + \beta_3 * \text{male}_i * \text{age}_{ij} + \beta_4 * \text{age}_{ij} * \text{age}_{ij} + \varepsilon_{ij} \quad i = 1..200 \quad j = 1..5$$

In this model, *male* is a dummy variable indicating the gender of i th child, *age* is the age of i th child at j th visit. The error terms ε_{ij} have both time and subject factors, so it could not be assumed uncorrelated. Without a good estimation for the variance of error term, the estimate of outcome and model fitting might be inaccurate. The matrix is a picture of correlations of errors for repeated measures (e.g. visit1 v.s visit2, visit2 v.s visit3.). GENMOD procedure allows specifying different structures of this covariate matrix of residuals and also estimates model using Generalized Estimating Equation (GEE) method for correlated outcome with repeated measurements.

Three types of covariate matrix for residuals are examined in this model: First, Compound symmetry, each outcome variable has unique variance and the correlation between them is constant; Second, Independence, there is no

correlation over repeated measures(visits) or subjects between the error term of each outcome; Third, First Order Autoregressive, correlations between error terms declines based on distance.

The following are parts of results and codes of GENMOD procedure:

Covariance Structure	GEE Fit Criteria(QIC)
Compound Symmetry	539.6898
Independent	523.3782
Autoregressive(1)	525.8675

Analysis Of GEE Parameter Estimates(Independent Structure)						
Empirical Standard Error Estimates						
Parameter	Estimate	Standard Error	95% Confidence Limits		Z	Pr > Z
Intercept	7.8301	1.0108	5.8491	9.8111	7.75	<.0001
male	0.2768	1.1029	-1.8848	2.4384	0.25	0.8018
AGE	-0.3132	0.0492	-0.4096	-0.2168	-6.37	<.0001
male*AGE	0.0045	0.0353	-0.0647	0.0737	0.13	0.8978
AGE*AGE	0.0017	0.0006	0.0006	0.0028	2.94	0.0033

```
Proc Genmod Data=nepalbin descending;
Class pid ;
Model bfstatus=male age age*male age*age / d=binomial link=logit;
Repeated subject=pid /type=ind corrw covb ;
run;
```

Quasilikelihood under the Independence model Criterion (QIC) is a statistic comparing models fit as GEE is not likelihood method, other criteria are not available. Based on QIC(lower is better), model with independence covariance structure fits this data best. The second table is model estimates from independence covariance structure. The result indicates that children's gender does not affect the probability of breastfeeding; when children's age increase, the probability of breastfeeding decreases, and the magnitude of this effect will diminish later.

Conclusion

GENMOD procedure gives an opportunity to construct logistic regression model for possible correlated outcome data with repeated measurements. The application of GENMOD for Nepal's children data confirmed that efforts are needed to encourage women in Nepal for breastfeeding practice as well as extend breastfeeding while children's ages increase. Future research might be focused on hypothesis tests for correlated outcome data and application in SAS system.

Reference:

Diggle PJ, Liang KY, Seger SL (1995). Analysis of Longitudinal Data. Oxford University Press, New York.

Francesca Dominici(2009). Longitudinal Logistic Regression: Breastfeeding of Nepalese Children,
http://www.biostat.jhsph.edu/~fdominic/teaching/bio655/labs/labs2008/Lab11_binary.nepal09.pdf.

SAS Institute Inc (2010). SAS/STAT(R) 9.2 User's Guide, Second Edition.

United Nations Children's Fund (2006). Nepal Demographic Health Survey.

SAS and all other SAS Institute Inc. product or service names are registered trademarks or trademarks of SAS Institute Inc. in the USA and other countries. ® indicates USA registration. □ □ Other brand and product names are trademarks of their respective companies.