Comparison of Marginal and Random Effects Models: The effect of age on renal function

McCrink, L., Marshall, A., & Cairns, K. (2010). Comparison of Marginal and Random Effects Models: The effect of age on renal function. 77-78. Paper presented at Annual Conference on Applied Statistics in Ireland (CASI), Portrush, United Kingdom.

Comparison of Marginal and Random Effects Models: The effect of age on renal function

Lisa McCrink, Adele Marshall and Karen Cairns
Centre for Statistical Science and Operational Research (CenSSOR)
Queen's University, Belfast, UK

June 2, 2010

Abstract

There are various approaches currently available to model longitudinal data [1]. The purpose of this research is to compare two such models which are extensions of generalized linear models (GLM): marginal and random effects models. These will be compared to determine which model is most suitable to handle measurements which are unbalanced and occur at uneven time intervals.

The application of interest is the fluctuations in a patient's kidney function, focusing on the effect of age on a patient's estimated Glomerular Filtration Rate (eGFR). The dataset considered comprises of the repeated measurements of patients over a two year period commencing in 2001. This analysis focused on 100 patients aged over 50 with more than one observation in the time period. The measurements are unbalanced and occur at uneven time intervals.

Marginal models were fitted using both Generalized Estimating Equations (GEE) and Quasi-Least Squares (QLS). GEE was modelled as it is the most commonly used method for estimation of the working correlation matrix. QLS, however, was also modelled as it has the ability to fit a Markov correlation matrix, which is desirable as it incorporates stronger correlations between responses that are closer together in time. GEE was fitted using SAS, R and SPSS. QLS was fitted using SAS, R and Matlab. Both of these estimation techniques

were compared using the Quasi-Information Criterion (QIC).

The random effects model was implemented using a multilevel mixed effects model (given that 59% of the variation between subjects can be accounted for due to the dependency of repeated measurements within individuals). Multilevel models were fitted using SAS, R and MLwiN and compared using the Akaike's Information Criterion (AIC).

The optimal marginal model was that fitted in SAS based on GEE using an exchangeable correlation matrix. With this it can be inferred that the average decline of the population's eGFR with age is approximately 0.6%, in agreement with renal literature [2]. None of the computer packages converged using the Markov correlation matrix for QLS estimation of the marginal model. The optimal multilevel model was fitted using an unstructured correlation matrix in MLwiN and also indicates that an individual's eGFR with age declines by approximately 0.8%. As the multilevel model was seen to be computationally less intensive, its use is recommended for modelling larger unbalanced datasets.

To conclude, the models discussed in this paper provide different interpretations of the regression parameters being estimated and thus their use is dependent on the purpose of the study. One limitation of both models is their inability to handle non-linear fluctuations in an individual's eGFR over time. Further research can include the use of a transitional model, as it can handle response fluctuations. Thus a comparison can then be made between the three commonly used GLMs to model longitudinal data.

References

[1] Diggle, P.J., Heagerty, P., Liang, K. & Zeger, S.L. 2002, *Analysis of Longitudinal Data*, Second Edition edn, Oxford University Press, United States.

[2] Lindeman, R.D., Tobin, J. & Shock, N.W. 1985, Longitudinal-Studies on the Rate of Decline in Renal-Function with Age, *Journal of the American Geriatrics Society*, vol. 33, no. 4, pp. 278-285.