

## **A COMPARISON OF STATISTICAL METHODS FOR THE ANALYSIS OF BINARY REPEATED MEASURES DATA WITH ADDITIONAL HIERARCHICAL STRUCTURE**

Elmabrok Masaoud<sup>1</sup>, Henrik Stryhn<sup>1</sup>

<sup>1</sup> Department of Health Management, Atlantic Veterinary College, University of Prince Edward Island, Charlottetown, Canada

The primary objective of the study was to compare statistical methods for the analysis of binary repeated measures data with an additional hierarchical level. In veterinary epidemiological research, such data are commonly encountered, and one motivating setting for the present study was records of high somatic cell counts in milk samples obtained by approximately monthly sampling throughout the lactations of cows in dairy herds. As the basis of a simulation study, random effects true models with autocorrelated ( $\rho = 1, 0.9$  or  $0.5$ ) subject random effects were used. In general, the settings of the simulation were chosen to reflect a real somatic cell count dataset, except that the within-cow time series were balanced, complete and of fixed length (4 or 8 time points). Among the estimation procedures considered were: Ordinary Logistic Regression (OLR), Alternating Logistic Regression (ALR), Generalized Estimating Equations (GEE), Marginal Quasi Likelihood (MQL), Penalized Quasi Likelihood (PQL), Pseudo Likelihood (REPL) (as implemented in PROC GLIMMIX in SAS), Maximum likelihood via numerical integration (ML) and Bayesian Markov chain Monte Carlo (MCMC). Results showed that the ALR estimation procedure performed well and attained a CI coverage close to nominal level. The REPL gave substantially biased estimates with large relative bias (generally greater than 10%). The OLR procedure failed to attain good CI coverage. For datasets generated by the random intercept model (i.e.,  $\rho = 1$ ), the ML, MCMC and PQL estimation procedures performed well.

A secondary objective of the study was to assess the impact of missing values in the data structures described above. Three different scenarios of simulated incomplete datasets were considered. The first scenario corresponded to a combination of three types of missingness patterns present in a somatic cell count dataset: “pre drop-outs” (time points within a subject prior to study onset were considered as missing values), drop-outs (subjects leave the study at some point in time), and intermittent missing values. The second and third scenarios involved only drop-outs, and corresponded to either: moderate or high percentage of missing values at random (MAR), or missing values not at random (MNAR), respectively. Diggle’s (1994) logistic random effects model was adapted to simulate the missing values. In the first scenario, all estimation procedures except OLR performed well and produced estimates with small relative bias (generally less than 5%) for levels of missingness that roughly corresponded to the real somatic cell count data. For random intercept models with drop-outs missing at random, ML and MCMC procedures performed well, while ALR worked well only at moderate percentages of missing values. However, PQL produced substantially biased estimates with large relative bias. In the simulations involving drop-outs missing not at random, ALR, ML and MCMC performed fairly well and produced estimates for effects constant over time with small relative bias. Finally, the comparison between ARL and MCMC procedures in datasets that contained both autocorrelation ( $\rho < 1$ ) and missing values was broadly consistent with the findings in datasets where only one of these were present.