

Methods for the evaluation of binary logistic regression predictions

**Caroline Fenlon**<sup>1</sup>, Luke O'Grady<sup>2</sup>, John Dunnion<sup>1</sup>, <sup>1</sup>University College Dublin, Dublin 4, Ireland;  
<sup>2</sup>University College Dublin, Dublin 4, Ireland. Contact: caroline.fenlon@ucdconnect.ie

**Purpose:**

Logistic regression is a statistical modelling technique that is widely used for the analysis of binary events in epidemiology. Similar to some forms of linear regression, logistic regression works by using Maximum Likelihood estimation to determine a function mapping one or more independent variables to a binary outcome. All major statistical software packages allow such models to be built with relative ease, but evaluation of the resultant predictions is not widely documented. Existing software facilitates the use of tests for determining overall goodness of fit (e.g. the likelihood ratio test for comparisons with the null model), but there is little focus on predictive capability.

**Methods:**

This study describes the application of visual and statistical methods for the evaluation of logistic regression, including models with random effects, using the statistical programming language R. The usage and suitability of each method for binomial and binary data is examined by application to dairy cattle conception risk.

**Results:**

Pearson residuals help to identify outlying records which are predicted poorly by the model. Calibration plots can be used for broad evaluation of grouped predictions, to identify overall bias or error. The Hosmer-Lemeshow test also groups predicted probabilities and uses a chi-square test to determine if the group predictions are statistically different from the real outcomes. Evaluation of predictions for discretised independent variables highlights error in a more fine-grained manner.

**Conclusions:**

In-depth analysis of individual predictor variable responses is useful for checking the suitability of the model for prediction at a more granular level.

**Relevance:**

The evaluation methods most prevalent in literature are those suitable for binomial data (where the outcome is coded as the proportion of events occurring within a group with identical values in each predictor variable), but which do not apply to binary data, where records may be impossible to group due to continuous independent variables (i.e. outcome is 1 or 0 for a single record). The methods described here are applicable to both binomial and binary data.