

# **A BOOTSTRAP APPROACH TO IDENTIFYING FACTORS AFFECTING WAITING TIME TO HEART TRANSPLANTATION**

R L Potter, P White  
UK Transplant/University of the West of England

Logistic regression is widely used to identify important risk factors for binary response variables. Many statistical packages have automated selection procedures; forward selection, backward elimination and stepwise selection. One way of validating a model obtained using automated selection procedures is to use a bootstrap technique as a variable selection tool. Bootstrapping uses random sampling of observations with replacement, so that an observation can appear several times or not at all.

A cohort consisting of 1008 patients listed for routine heart transplant in the UK, between 1 January 1999 and 31 December 2004, was used to investigate three methods of bootstrapping. The response variable of interest is whether or not a patient received a transplant, and risk factors will include patient demographic and clinical factors.

Three bootstrapping methods were used to validate the original automated selection models. The first method uses a cut-off point to decide which variables are significant and should be included in a model. The second uses a random probe variable to decide which variables should be included. The third method uses a series of fake variables to select the variables in the model.

The bootstrapping procedure was repeated 200 times on the full data set to give 200 samples, with each sample having the same number of observations as the original data set. For each of the 200 bootstrap samples for the three bootstrap methods and the three variable selection procedures, logistic regression models were analysed and the number of times a variable was selected was recorded.

A SAS<sup>TM</sup> macro has been developed to automate the bootstrapping process. The macro automates the results, which are in the format of graphs, and exports them directly to Microsoft Word<sup>TM</sup>.