Model-Building Approach
in Multiple Binary Logit Model for Coronary Heart Disease

[1]H. J. Zainodin and [2]G. Khuneswari

School of Science and Technology
Universiti Malaysia Sabah, Locked Bag No. 2073,
88999 Kota Kinabalu, Sabah, Malaysia
E-mail: [1]zainodin@gmail.com, [2]khuneswari@gmail.com

ABSTRACT

This paper develops a procedure to find the best model. An illustration of the model-building approach in "Multiple Binary Logit" analysis has been introduced. The dependent variable of Multiple Binary Logit model is a qualitative nature (taking a value of 0 or 1). Besides introducing multiple single independent variables into the model, all possible combinations of generated interaction variables are included in the model. In order to obtain a set of selected models, a progressive elimination (one by one, least significant first) of the insignificant variables is employed. It was also proposed to use the modified Eight Selection Criteria (8SC) by replacing SSE (sum square of error) Deviance Statistic (G^2) to finally single out the best model. A numerical illustration (case study) on coronary heart disease (CHD) was included in order to get a clear picture of the procedure of getting the best Multiple Binary Logit model. In addition, there are three quantitative (age, cholesterol level and body-mass index (BMI)) and five qualitative independent variables (4 blood pressure categories and smoking habit). Detailed procedure is exposed, illustrated and explained using this case study. The model building approach through the Multiple Binary Logit model was established. It was found that interaction variables as age and BMI, BMI and cholesterol level, age and blood pressure category, SBP 100-129 and DBP 60-79, age and blood pressure category, SBP 130-139 and DBP 80-89, BMI and blood pressure category, SBP 130-139 and DBP 80-89, BMI and blood pressure category, SBP >140 and DBP >90 and cholesterol level and blood pressure category, SBP 100-129 and DBP 60-79 are significant in the best model obtained. The dummy variables blood pressure category, SBP >140 and DBP >90 and smoking/non smoking are significant in the best model obtained.

Keywords: model-building approach, Multiple Binary Logit, interaction variables, elimination procedure, M8SC, Deviance Statistics