Model-building on survivability of upper gastrointestinal bleed patient's

To cite this article: Khuneshwari Gopal Pillay and Siti Aisyah Mohd Padzil 2018 J. Phys.: Conf. Ser. 1132 012067

View the article online for updates and enhancements.
Model-building on survivability of upper gastrointestinal bleed patient’s

Khuneshwari Gopal Pillay1,a and Siti Aisyah Mohd Padzil1

1Universiti Tun Hussein Onn Malaysia Department of Mathematics and Statistics, Faculty of Applied Science and Technology, Universiti Tun Hussein Onn Malaysia, Pagoh Kampus, KM 1 Panchor Road, 84000, Muar Johor Malaysia

E-mail: akhuneswari@uthm.edu.my

Abstract. Statistical modelling by using Multiple Binary Logit (MBL) or Logistic Regression (LR) on a medical data has been a common practice by the researchers. However, there are no agreed guidelines for how best to carry out model-building using MBL to obtain the best model. This research will propose an appropriate guideline for beginners and demonstrate the flow of model-building process using Rockall score data as well as highlighting the significant factors of survivability of Upper Gastrointestinal bleed (UGIB) patients in Sabah. Rockall score is a scoring system used in identifying the risk of survivability for patients with UGIB. The patient’s data were retrieved from Hospital Queen Elizabeth in Sabah. Seven categorical variables related to the Rockall scoring system were studied and the steps to obtain best model using MBL were illustrated in four phases. The phases include all possible models, selected models, best model and goodness-of-fit test. All possible models were considered without interaction variables. A progressive elimination (one by one, least significant first) of the insignificant variables is carried out to a set of selected models (with significant variables). Model selection criteria AIC, corrected AIC (AICc) and BIC were used to single out the best model among the selected models. Pearson chi-square test and deviance chi-square test were carried out to ensure the best model validity and appropriateness. The results showed that the factors affecting the survivability of UGIB patients in Sabah are shock score, comorbidity and rebleed. In conclusion, the study showed that the Rockall scoring system had satisfactory validity for the prediction of shock score, comorbidity and rebleeding in patients with UGIB. There was a negative relationship between the clinical Rockall scores and patient outcomes in terms of shock score and comorbidity.

1. Introduction

Model-building approach is one of the key areas of interest in the application of statistical modeling. The statistical modeling methods will be chosen based on the dependent and independent variables. Multiple Binary Logit (MBL) is a regression model where the dependent variable is binary. When the dependent variable has two possible qualitative outcomes and represented by a binary indicator variable taking on the values 0 and 1, it is called a binary logit model. According to Zainodin and Khuneshwari [1], the MBL is the extension of logit model and known as qualitative choice model. It is used for prediction of the probability of occurrence of an event. Maximum likelihood (ML) is used for parameter estimation of multiple binary logit model.

This research illustrates the model-building approach using Rockall Score data from UGIB patients, and highlight the most significant factors of mortality for UGIB patients. The whole procedures of
obtaining the best MBL model will be explained step by step to provide a clear guideline of model-
building by using MBL model.

2. Methodology

2.1. Multiple Binary Logit (MBL)
MBL model or often called as Logistic Regression model is a form of regression with binary dependent
variable. In this research, the outcome of the study is the patient’s survivability which takes on value
1(alive) and 0 (dead). The general MBL model as suggested by Kutner et al. [2] is

\[
Y_i = \beta_0 + \beta_1 X_{1i} + \beta_2 X_{2i} + \ldots + \beta_q X_{qi} + u_i \tag{1}
\]

\[
Y_i = \ln \left( \frac{P_i}{1-P_i} \right) \tag{2}
\]

and

\[
P_i = \frac{\exp(\beta_0 + \beta_1 X_{1i} + \beta_2 X_{2i} + \ldots + \beta_q X_{qi})}{1 + \exp(\beta_0 + \beta_1 X_{1i} + \beta_2 X_{2i} + \ldots + \beta_q X_{qi})} \tag{3}
\]

where the binary dependent variable is denoted by \( Y \), \( X_j \) is the \( j^{th} \) independent variable where \( j = 1,2 \ldots q \), the constant term of the model is denoted by \( \beta_0 \), \( \beta_j \) is the \( j^{th} \) coefficient of independent variable where \( j = 1,2 \ldots q \), \( u \) is the random error of the model and \( P_i \) is the probability of event occurs.

2.2. Model-building phases
According to Zainodin and Khuneswari [1], the process of model-building for multiple regression model
involves four phases. Below is the proposed model-building phases of MBL model.

| PHASE 1: All Possible Models |
|-----------------------------|
| Single independent variables and all possible product of related single independent variables (interaction variables) |

| PHASE 2: Selected Models |
|--------------------------|
| Elimination, discard insignificant variables |

| PHASE 3: Best Model |
|---------------------|
| Using AIC, AICc and BIC |

| PHASE 4: Goodness-of-fit Test |
|------------------------------|
| Deviance and Pearson Chi Square Test |

Figure 1. Four Phases of Model-Building Approach using Multiple Binary Logit

2.2.1. All possible models. The first phase of building a model is to list out all possible combinations of
variables to form a model. According to Zainodin and Khuneswari [1], all possible models with no
interaction variables can be calculated using Equation (4).

\[ N = \sum_{j=1}^{q} \binom{n}{c_j} \]
where,
\[ N = \text{number of all possible models} \]
\[ q = \text{number of single independent variables } j = 1, 2, 3 \ldots, q \]

2.2.2. Selected Models. In the second phase, the coefficient for the models were estimated using maximum likelihood (ML) estimation and several tests were conducted to obtain selected models (models with only significant variables). Coefficient test and Wald test will be carried out to select and eliminate insignificant variables from all possible models. Hosmer and Lemeshow [4] stated that Wald test and Likelihood Ratio test is the same. Coefficient test helps to eliminate any insignificant (non-contributing) variables while the Wald test were conducted to justify the elimination of insignificant variables in the model, so that only significant variables that related to the dependent variables remain in the final model.

Coefficient test is needed to identify and eliminate the insignificant variables. The elimination of insignificant variables is done by comparing the p-value for each variable in the model. Variable with p-value more than 0.05 indicate the insignificant variable. The processes only allow one variable in the model to be eliminated in a single run. Therefore, variable with highest p-value and more than 0.05 is eliminated first and the new model is rerun again. The processes of omitting variable continue until all insignificant variables are eliminated.

Wald test will justify the removal of insignificant variables, Khuneswari et al. [3]. Hypothesis testing for the Wald test includes Restricted model (R) and Unrestricted model (U). Restricted model is the selected model, which was obtained after omitting the insignificant variables in coefficient test, whereas the Unrestricted model is the initial possible model. Below is the situation for Unrestricted model and Restricted

\[ \text{U: } Y = \beta_0 + \beta_1 X_{1i} + \beta_2 X_{2i} + \beta_m X_{mi} + \beta_{m+1} X_{(m+1)i} + \ldots + \beta_q X_{qi} + u_i \]  
\[ \text{R: } Y = \beta_0 + \beta_1 X_{1i} + \beta_2 X_{2i} + \ldots + \beta_q X_{qi} + u_i \]

The Restricted model was obtained by removing \( X_{m+1}, X_{m+2}, \ldots, X_q \) (not necessarily in this order) from Unrestricted model. Hence, the hypothesis of the Wald test for removing the variables is

\[ H_0 : \text{Restricted Model is true} \]
\[ H_1 : \text{Unrestricted model is true} \]

According to Hosmer and Lemeshow [4], testing for the significance of the coefficient in MBL model compares the observed value of the response variable to predicted values. The comparisons are based on log-likelihood function as follows:

\[ G = -2[LL(R) - LL(U)] \]  

where,
\[ G = \text{deviance that follow chi-square distribution} \]
\[ LL(R) = \text{log-likelihood of restricted model} \]
\[ LL(U) = \text{log-likelihood of unrestricted model} \]

Since the statistic \( G \) value follows the chi-square distribution, \( \chi^2(V) \) denote chi-square random variable with \( V \) degree of freedom. Hosmer and Lemeshow [4] stated that the \( p \)-value associated with the test can be obtained by using chi-square distribution table using following information

\[ P[\chi^2(V)] > G \]

\[ V = V_R - V_U, \text{ note that } V_R > V_U \]

where
\[ G = \text{deviance} \]
\( V \) = degree of freedom  
\( V_R \) = degree of freedom for smaller model (restricted)  
\( V_u \) = degree of freedom for initial model (unrestricted)

\( V \) is obtained from the subtraction between restricted model and unrestricted model degrees of freedom. Alternatively, \( V \) is the number of variables omitted from the unrestricted model. The decision is accept \( H_0 \) if the \( p \)-value is larger than 0.05. Therefore, elimination of insignificant variable in a model will be justified. The test was then carried out for each remaining selected models.

### 2.2.3. Best Model

In order to obtain the best model, several model selection criteria will be calculated for each selected model. The model with the minimum value for all criteria is chosen as the best model. The model selection criteria, Akaike Information Criterion (AIC) by Akaike [5], corrected AIC (AICC) by Hurvich and Tsai [6] and the Bayesian Information Criterion (BIC) by Schwarz [7] will be used in this research. The model selection criteria are

\[
AIC = -2 \log L(M) - 2p \\
AIC_c = -2 \log L(M) - 2(q + 1) \frac{n}{n-q-2} \\
BIC = -2 \log L(M) - (q + 1) \log(n)
\]

where,

- \( L(M) \) = minimum value for likelihood function of model \( M \),
- \( n \) = number of observations
- \( q + 1 \) = number of parameters.

### 2.2.4. Goodness-of-fit-test

After identifying the best model, goodness-of-fit test will be applied to ensure the model validity and appropriateness. There are two tests suggested by Kutner et al. [3] which are Pearson Chi-Square goodness-of-fit test and Deviance goodness-of-fit test. The first test which is Pearson residual is defined Kutner et al. [2] as follows:

\[
r_{pi} = \frac{Y_i - \hat{P}_i}{\sqrt{\hat{P}_i(1-\hat{P}_i)}}
\]

where,

- \( \hat{P}_i \) = estimated probability
- \( Y_i - \hat{P}_i \) = ordinary residual,
- \( \sqrt{\hat{P}_i(1-\hat{P}_i)} \) = estimated standard error of \( Y_i \) for \( i = 1,2,\ldots,n \),
- \( r_{pi} \) = Pearson residual

Khunswari et al. [3] suggested the steps of Pearson Chi-Square goodness-of-fit. The steps are as follows:

**Step 1:** Hypothesis Testing  
\( H_0: E(Y) = [1 + \exp(-X^T\beta)]^{-1} \)  
\( H_1: E(Y) \neq [1 + \exp(-X^T\beta)]^{-1} \)

**Step 2:** Test Statistic  
\[ X_{r_{pi}}^2 = \frac{\sum_{i=1}^{n} \left(Y_i - \hat{P}_i\right)^2}{\sqrt{\hat{P}_i(1-\hat{P}_i)}} \]  

**Step 3:** Set \( \alpha = 5\% \). The critical value for Pearson chi-square goodness-of-fit is  
\[ X^2_{critical} = X^2_{(1-\alpha,n-k-1)} \]

**Step 4:** If \( X_{r_{pi}}^2 \leq X^2_{critical} \), then accept null hypothesis. Otherwise, accept the alternative hypothesis. The model is valid or appropriate when the null hypothesis is
accepted.

The second test is Deviance residual test is defined by Kutner et al. [2] as follows:

\[
dev_i = \text{sign} (Y_i - \hat{P}_i) \sqrt{-2[Y_i(\ln \hat{P}_i) + (1 - Y_i) \ln (1 - \hat{P}_i)]}
\]

where

\[
sign = \begin{cases} + \text{ when } Y_i \geq \hat{P}_i, & \text{for } i = 1, 2, \ldots, n \\ - \text{ when } Y_i < \hat{P}_i \\
\end{cases}
\]

The steps of Deviance goodness-of-fit test suggested by Khuneswari et al. [3] is as follows:

**Step 1:** Hypothesis Testing
\[ H_0: E(Y) = [1 + \exp(-X^T \beta)]^{-1} \]
\[ H_1: E(Y) \neq [1 + \exp(-X^T \beta)]^{-1} \]

**Step 2:** Test Statistic
\[ G^2 = -2 \sum_{i=1}^{n} [Y_i(\ln \hat{P}_i) + (1 - Y_i) \ln (1 - \hat{P}_i)] \]

**Step 3:** Set \( \alpha = 5\% \). The critical value for Deviance goodness-of-fit is
\[ \chi^2_{\text{critical}} = \chi^2_{1-\alpha,n-k-1} \]

**Step 4:** If \( G^2 \leq \chi^2_{\text{critical}} \), then accept null hypothesis. Otherwise, accept the alternative hypothesis. The model is valid or appropriate when the null hypothesis is accepted.

To support the results from goodness-of-fit tests, scatter plot of residuals can be used to check the model appropriateness. Kutner et al. [2] mentioned that there are three common residual plots used in the Logistic Regression analysis, which are ordinary residual against estimated probability, Pearson residual against estimated probability and Deviance residual against estimated probability. If the model is appropriate, the plot should approximately result in horizontal line with zero intercept.

### 3. Modelling of Rockall Score Data

#### 3.1. Rockall score data

The Rockall scoring system is a method that is used to determine the risk of death for patients with UGIB. By knowing the risk of death of a patient, the hospital will be able to provide better care for patients with high risk of mortality. According to Vreeburg et al. [8], Rockall scoring systems helps in identifying low risk patients for early discharge as well as selecting high risk patients for intensive treatment. Data of Rockall Score for UGIB consist of 410 samples and comprise of seven qualitative independent variables which are age score, shock score, comorbidity, diagnosis score, major score, rebleed and Rockall group. This data was retrieved from QEH in Sabah.

Table 1 shows the data descriptions for each variable in the Rockall score dataset. The data comprises of qualitative data, hence the process of model building using only qualitative data without interaction will be explored. The dependent variable represents the survival/mortality of the patients. In the data set, the survivability of the patients \( Y \) is represented as 1 if the patients survive and 0 if otherwise. Patient’s Age score \( X_1 \) takes on the value 0, 1 and 2 which are <60 years old, between 60-79 years old and age \geq 80 years old respectively. For the second variable, shock or circulatory shock \( X_2 \) is when an organ or body tissue does not receive a normal blood flow. In the dataset, no shock means that the blood flow is normal, tachycardia means abnormal rapid heart rate and hypertension is a case when there is abnormal high blood pressure.
Table 1. Data description.

| Variable | Description |
|----------|-------------|
| \( Y \)  | Survival of Patients  
1 if the patient survives  
0 if the patient not survives |
| \( X_1 \) | Age Score  
0: if age <60  
1: if age 60-79  
2: if age ≥80 |
| \( X_2 \) | Shock Score  
0: No shock  
1: Tachycardia  
2: Hypotension |
| \( X_3 \) | Comorbidity  
0: Nil major  
1: Cardiac failure, IHD, others  
2: Renal failure, liver failure, disseminated Malignancy |
| \( X_4 \) | Diagnosis Score  
0: Mallory-Weiss Tear, No Lesion  
1: All other diagnosis  
2: Malignancy of UGI |
| \( X_5 \) | Major Score  
0: None or Dark Spots  
1: Blood in Upper GIT, Adherent Clot, Visible Spurting Vessel |
| \( X_6 \) | Rebleed  
1: Yes  
2: No |
| \( X_7 \) | Rockall Group  
1: Low Risk  
2: Medium Risk  
3: High Risk |

The next variable which is co-morbidity \((X_3)\) is defined as the presence of two or more chronic diseases or conditions in a patient. Comorbidity is based on reference standard diagnostic criteria, including cardiovascular and cerebrovascular disease, chronic obstructive pulmonary disease, chronic liver disease and cancer. Nil major which is scored as 0 shows patient with no major disease. Score 1 in the co-morbidity score shows present of cardiac failure, IHD (Ischemic Heart Disease) and other disease in patients. IHD is one of cardiovascular disease. Lastly, for Score 2, it shows that the patients have renal failure (renal is associated with kidney), liver failure and disseminated malignancy. Disseminated malignancy is a case where the cancer was extremely spreading up to stage 4.

Mallory Weis Tear is scored as 1 in diagnosis score \((X_4)\). The term means that the patient’s esophagus is teared due to too much vomit. The fifth predictor which is a major score \((X_5)\) is the scoring for UGIB. It is scored as 0 if there is no bleeding or dark spot (clean ulcers) and is scored as 2 if it there is bleeding in UGIB. The sixth variable which is rebleed \((X_6)\) score is scored as 0 if no rebleed and score 1 if rebleed happen. For the last independent variable which is a Rockall group \((X_7)\), it is based on patients Rockall score. Table 2 explains the Rockall group form from using Rockall scores.
Table 2. Rockall group from Rockall scores.

| Rockall Score | Rockall Group | Description |
|---------------|---------------|-------------|
| 0-2           | 1             | Low risk    |
| 3-7           | 2             | Medium Risk |
| 8-11          | 3             | High Risk   |

3.2. Model-building phases

3.2.1 Stage 1: All Possible Models. The process of obtaining the best model starts by listing all possible models. Since model-building of Rockall score dataset only considers individual variables, the number of all possible models were calculated using Equation (4):

\[ N = 1(7C_1) + 1(7C_2) + 1(7C_3) + 1(7C_4) + 1(7C_5) + 1(7C_6) + 1(7C_7) \]

\[ = 127 \text{ possible models} \]

Table 3. All possible models.

| Number of variables | Number of Models | Model        |
|---------------------|-----------------|--------------|
| 1                   | 7 models        | M1-M7        |
| 2                   | 21 models       | M8-M28       |
| 3                   | 35 models       | M29-M63      |
| 4                   | 35 models       | M64-M98      |
| 5                   | 21 models       | M99-M119     |
| 6                   | 7 models        | M120-M126    |
| 7                   | 1 model         | M127         |
| Total               | 127 models      |              |

Table 3 shows all possible models will be used for model-building of Rockall datasets. There are 127 possible models.

3.2.2 Stage 2: Selected Models. All variables in 127 possible models were tested one by one to eliminate any insignificant variable and to ensure that only significant variables were left in the model. Two tests were involved in the process of identifying selected model. The tests were Coefficient Test and Wald Test. For clear understanding, the process of obtaining selected model is shows step by step by using model M120. The model M120 is:

\[ M120: Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 + \beta_4 X_4 + \beta_5 X_5 + \beta_6 X_6 + u \]

Table 4. Insignificant variables elimination for model M120.

| Variables in model M120 | \( p \)-value | M120.1 | M120.2 | M120.3 | M120.4 |
|-------------------------|---------------|--------|--------|--------|--------|
| \( X_1 \)               |               | 0.5309 |        |        |        |
| \( X_2 \)               |               | 0.0678 | 0.0650 | 0.0680 | 0.0427 |
| \( X_3 \)               |               | 0.0013 | 0.0014 | 0.0016 | 0.0012 |
| \( X_4 \)               | \textbf{0.7056} |        |        |        |        |
| \( X_5 \)               |               | 0.2534 | 0.2548 |        | \textbf{0.2601} |
| \( X_6 \)               |               | 0.0108 | 0.0084 | 0.0085 | 0.0031 |

Table 4 shows the process of eliminating insignificant variables. The variable \( X_4 \) was eliminated in the first run (M120.1) followed by \( X_1 \) and \( X_5 \) in the second run (M120.2) and third run (M120.2) respectively. The variables were eliminated as it has the highest \( p \)-value and more than 0.05 for its
corresponding run. Since the fourth run shows all variables in model M120.4 were significant, the process of omitting insignificant variable stopped. Note that the model M120.4 is the same as M46 in all possible models list.

M120.4 or M46:  
\[ \hat{Y} = \hat{\beta}_0 + \hat{\beta}_2X_2 + \hat{\beta}_3X_3 + \hat{\beta}_6X_6 \]

The Unrestricted model and Restricted Model for Wald test of model M120 are:

- **U**:  
  \[ Y = \beta_0 + \beta_1X_1 + \beta_2X_2 + \beta_3X_3 + \beta_4X_4 + \beta_5X_5 + \beta_6X_6 \]

- **R**:  
  \[ Y = \beta_0 + \beta_2X_2 + \beta_3X_3 + \beta_6X_6 \]

Hypothesis for Wald Test:

- **H₀**: Restricted model is true
- **H₁**: Unrestricted model is true

**Table 5.** Log-likelihood value for unrestricted model (U) and restricted model (R).

| Model           | Log-likelihood value |
|-----------------|----------------------|
| Unrestricted    | 165.7953             |
|Restricted Model | 165.0049             |

Table 5 show the value of likelihood for both models. The value of \( V \) is 3 because there are three variables were eliminated. The statistic \( G \) is

\[ G = -2[(165.0049) - (165.7953)] = 1.5808 \]

Therefore the \( P[\chi^2(3)] > 1.5808 \] = 0.0722. Since the \( p \)-value is larger than 0.05, the \( H_0 \) is accepted, hence the removal of insignificant variables in the model is justified. Model M46 is listed in the selected models. Table 6 shows all the listed selected models for Rockall score data.

**Table 6.** Selected models for Rockall score data.

| Model | Selected Models |
|-------|-----------------|
| M2    | \( \hat{Y} = \hat{\beta}_0 + \hat{\beta}_2X_2 \) |
| M3    | \( \hat{Y} = \hat{\beta}_0 + \hat{\beta}_3X_3 \) |
| M5    | \( \hat{Y} = \hat{\beta}_0 + \hat{\beta}_6X_6 \) |
| M6    | \( \hat{Y} = \hat{\beta}_0 + \hat{\beta}_6X_6 \) |
| M7    | \( \hat{Y} = \hat{\beta}_0 + \hat{\beta}_7X_7 \) |
| M14   | \( \hat{Y} = \hat{\beta}_0 + \hat{\beta}_2X_2 + \hat{\beta}_3X_3 \) |
| M17   | \( \hat{Y} = \hat{\beta}_0 + \hat{\beta}_2X_2 + \hat{\beta}_6X_6 \) |
| M21   | \( \hat{Y} = \hat{\beta}_0 + \hat{\beta}_3X_3 + \hat{\beta}_6X_6 \) |
| M28   | \( \hat{Y} = \hat{\beta}_0 + \hat{\beta}_6X_6 + \hat{\beta}_7X_7 \) |
| M46   | \( \hat{Y} = \hat{\beta}_0 + \hat{\beta}_2X_2 + \hat{\beta}_3X_3 + \hat{\beta}_6X_6 \) |
| M59   | \( \hat{Y} = \hat{\beta}_0 + \hat{\beta}_2X_2 + \hat{\beta}_3X_3 + \hat{\beta}_6X_6 + \hat{\beta}_7X_7 \) |

3.2.3. **Stage 3: Best Model.** Phase 3 explains the process in selecting the best model of Rockall Score from the listed selected models. A model that has the minimum value of model selection criteria (AICc, AIC, and BIC) and maximum value of the log likelihood indicates the best model. Table 7 shows the summary of model selection criteria for all selected models of Rockall data. Model M46: \( \hat{Y} = \hat{\beta}_0 + \hat{\beta}_2X_2 + \hat{\beta}_3X_3 + \hat{\beta}_6X_6 \) is selected as the best model since the value of log-likelihood is maximum and all other model selection criteria are minimum.
The hypothesis for model M46 is as follows:

\[ H_0: E(Y) = [1 + \exp(-X^T \beta)]^{-1} \]
\[ H_1: E(Y) \neq [1 + \exp(-X^T \beta)]^{-1} \]

The value of test statistics is \( \chi^2 = 258.3356 \) and for \( \chi^2_{critical} \) is \( \chi^2_{(0.95,402)} = 348.347 \). Since \( 258.3356 < 348.347 \), \( H_0 \) is accepted. Therefore, the best model M46 is appropriate and valid. The Deviance goodness-of-fit test, The hypothesis testing for best model is as follows:

\[ H_0: E(Y) = [1 + \exp(-X^T \beta)]^{-1} \]
\[ H_1: E(Y) \neq [1 + \exp(-X^T \beta)]^{-1} \]

The \( G^2 = 86.73909 \) and for \( \chi^2_{critical} \) the value was \( \chi^2_{(0.95,402)} = 348.347 \). Since \( 86.73909 < 348.347 \), \( H_0 \) is accepted. Therefore, it can be concluded that best model M46 is appropriate. Residual plots for ordinary residual against estimated probability, Pearson residual against estimated probability and deviance residual against estimated probability were shown in Figure 2, 3, and 4 respectively.

### Table 7. Model Selection Criteria; LL, AICc, AIC and BIC comparisons for selected models.

| Model | Possible Model | LL     | AICc   | AIC   | BIC   |
|-------|----------------|--------|--------|-------|-------|
| M2    | \( \hat{\gamma} = \hat{\beta}_0 + \beta_2 x_2 \) | 151.890 | -297.720 | -297.780 | -285.731 |
| M3    | \( \hat{\gamma} = \hat{\beta}_0 + \beta_3 x_3 \) | 156.020 | -305.9812 | -306.0403 | -293.9919 |
| M5    | \( \hat{\gamma} = \hat{\beta}_0 + \beta_5 x_5 \) | 150.6561 | -295.2531 | -295.3123 | -283.2638 |
| M6    | \( \hat{\gamma} = \hat{\beta}_0 + \beta_6 x_6 \) | 157.5817 | -307.0646 | -307.1634 | -291.0987 |
| M7    | \( \hat{\gamma} = \hat{\beta}_0 + \beta_7 x_7 \) | 162.3498 | -316.6008 | -316.6995 | -300.6349 |
| M14   | \( \hat{\gamma} = \hat{\beta}_0 + \beta_6 x_6 + \beta_7 x_7 \) | 159.8675 | -311.6363 | -311.7351 | -295.6704 |
| M17   | \( \hat{\gamma} = \hat{\beta}_0 + \beta_2 x_2 + \beta_3 x_3 \) | 154.1157 | -302.1722 | -302.2313 | -290.1828 |
| M21   | \( \hat{\gamma} = \hat{\beta}_0 + \beta_2 x_2 + \beta_6 x_6 \) | 155.2219 | -304.3846 | -304.4437 | -292.3952 |
| M28   | \( \hat{\gamma} = \hat{\beta}_0 + \beta_3 x_3 + \beta_6 x_6 \) | 158.8717 | -309.6446 | -309.7433 | -293.6787 |
| M46   | \( \hat{\gamma} = \hat{\beta}_0 + \beta_2 x_2 + \beta_3 x_3 + \beta_6 x_6 \) | 165.0049 | -319.8614 | -320.0099 | -299.9291 |
| M59   | \( \hat{\gamma} = \hat{\beta}_0 + \beta_3 x_3 + \beta_6 x_6 + \beta_7 x_7 \) | 160.1762 | -310.2039 | -310.3524 | -290.2716 |

#### 3.2.4. Stage 4: Goodness-of-fit test. The final process in model-building approach is goodness-of-fit test. The hypothesis for model M46 is as follows:

![Residuals](image1.png)  
**Figure 2.** Residuals for M46.

![Pearson Residuals](image2.png)  
**Figure 3.** Pearson Residuals for M46.
The ordinary residual plots above support both of goodness of fit test as it comes out with horizontal line and zero intercept. Pearson residuals and Deviance residuals do not follow the horizontal line but the plots do have zero intercept. According to Noraini et al. (2013), this may be due to small variance in the dependent variable. Despite from the residual plots do not follow the horizontal line trends, model M46 is still viewed as an appropriate and best model for Rockall Scoring data as (Kutner et al., 2008) stated that the plots are only used to justify the shape of regression curve and are not meant to analyze regression relationship. Therefore, M46 is concluded as best model for Rockall Score dataset. Model M46 is

$$\hat{Y} = 0.464 - 0.837X_2 - 0.808X_3 + 2.361X_6$$

The best model M46 showed that when variables $X_2$, $X_3$ and $X_6$ are 0, the probability of UGIB patient’s survivability is

$$P_i = \frac{exp^{0.464}}{1 + exp^{0.464}} = 0.6140 \approx 0.61$$

If there is one unit increase in shock score ($X_2$), the $\hat{Y}$ will decrease by 0.837. Therefore, the probability of UGIB patient's survivability will decrease also. Similarly, the probability of UGIB patient's survivability will decrease if there is increase comorbidity ($X_3$). The probability of UGIB patient's survivability will increase by 2.361 if there is a rebleed ($X_6$). Whereas, the probability of UGIB patient's survivability will increase by $2 \times 2.361 = 4.722$ if there is no rebleed.

4. Discussion and Conclusion

The analysis of model-building using MBL on Rockall data sets concluded that the best and the most appropriate model of Rockall is M46. It indicates all independent variables in M46 are significant. In other words, the independent variables in model M46 are the most influential risk factors in Rockall score data which will affect the survivability of UGIB patients. There were three factors that were discovered to associate with patient’s survivability using Rockall score data. Shock score ($X_2$), comorbidity ($X_3$) and rebleed ($X_6$) were found to be the most important factor of mortality for patients with UGIB in Sabah.

Shock score is a scoring made for abnormality in blood circulation which hence result in hypertension and tachycardia. Research conducted by Noraini et al. [9] and Chiu et al. [10] also conclude that Rockall scores show medical shock to be one of the risk factor of mortality. Shock is one of the critical condition in medical emergency and one of the most common causes of death for critical patients Tisherman et al. [11].

Comorbidity is a term use for multiple chronic diseases occurred in patients. For an example, usually patients with heart disease often have one or two comorbid diseases which are high blood pressure and diabetes. According to Valderas et al. [12], these multiple disease leads to a more complex clinical treatment and hence increased health care costs. Family doctors are faced with challenges in making
decisions about the treatment for patients with comorbidity because most of the clinical studies do not consider patients with multiple disease Osmun et al. [13]. Research conducted by Chiu et al. [10] concluded that multiple comorbidity is an important factor to predict mortality. Similarly, this research also concluded that the presence of comorbid disease is found to be an important risk factor of mortality among patients with UGIB in Sabah.

Recurrence of bleeding is one of the most important factors affecting prognosis, and early prediction and treatment of rebleeding would improve the outcome in patients with UGIB, as rebleeding is associated with high mortality. Rockall et al. [13] stated that rebleed or hemorrhage is described as the most influential factor of mortality for UGIB patients. Similarly, Vreeburg et al. [8] also point out that rebleeding is an important factor of mortality and occurs in 10–30% of successfully treated patients. Research conducted by Noraini et al. [9] and Chiu et al. [10] also concluded that rebleeding is an influential factor of mortality. Even though there are advance treatments available in treating patients with UGIB, rebleed still remains the life-threatening factor of UGIB patients Jairath [14].

In conclusion, the factors Shock score ($X_2$), comorbidity ($X_3$) and rebleed ($X_6$) were found to be the most important factors of mortality for patients with UGIB in Sabah. These factors provide a good indication on the mortality/survivability of UGIB patients in general. The research on upper gastrointestinal bleeding (UGIB) is common, costly, and potentially life-threatening and requires prompt assessment and aggressive medical management. Although the model-building process with 4 phases were clearly illustrated in the context of UGIB data set, however, the model selection as discussed in this research is well known for introducing additional uncertainty in the model-building process. The properties of the standard parameter estimates obtained from the selected model do not reflect the stochastic nature of the model selection process. Model averaging is proposed as an alternative to model selection, which is intended to overcome the underestimation of standard errors that is a consequence of model selection. Therefore, model averaging can be explored in the context of model-building using MBL model, and a comparison can be carried out between model selection and model averaging in the context of UGIB dataset to identify the significant factors affects the mortality.

References
[1] Zainodin H J and Khuneswari G P 2010 Model-Building Approach in Multiple Binary Logit Model using Coronary Heart Disease (Malaysian Journal of Mathematical Sciences vol 4) pp 107-133
[2] Noraini A Zainodin H J and Rick L B 2013 Risk factor determination on UGIB patients in Kota Kinabalu, Sabah Malaysia (Medical Sciences vol 13) pp 526-536
[3] Khuneswari G P Zainodin H J and Ho C M 2009 The Model Building Approach in Best Multiple Binary Logit Model with Higher Order Interaction Variables (Universiti Malaysia Sabah Master’s Thesis)
[4] Hosmer D W and Lemeshow S 2000 Applied Logistic Regression ed 2 (John Wiley and Sons Inc United States)
[5] Akaike H A 1978 Bhayesian analysis of the minimum aic procedure (Annal Of The Statistical Mathematics vol 30) pp 9-14
[6] Hurvich C M and Tsai C L 1991 Bias of the corrected aic criterion for underfitted regression and time series models (Biometrika vol 78) pp 499-509
[7] Schwarz G 1978 Estimating the dimension of a model (Annals of Statistics vol 6) pp 461–464
[8] Vreeburg E M Terwee C B Snel P Rauws E A J Bartelsman, J F W M Meulen J H P and Tytgat G N J 1999 Validation of the Rockall risk scoring system in upper gastrointestinal bleeding (Gut vol 44) pp 331-335
[9] Noraini A Zainodin H J and Rick L B 2013 Risk factor determination on UGIB patients in Kota Kinabalu, Sabah Malaysia (Medical Sciences vol 13) pp 526-536
[10] Chiu P W Y Ng E K W Cheung F K Y Chan F K L Leung W K Wu J C Y Wong V W S Yung M Y Tsoi K Lau J Y W Sung J J Y and Chung S S C 2009 Predicting mortality in patients
with bleeding peptic ulcers after therapeutic endoscopy (Clinical Gastroenterology and Hepatology vol 7) pp 311–316

[11] Tisherman S A Schmicker R H and Schreiber M A 2016 Detailed description of all death in both the shock and traumatic brain injury hypertonic saline trials of the resuscitation outcomes consortium (Annals of Surgery vol. 261) pp 586-590

[12] Valderas J M Starfield S Sibbald B Salisbury C and Rolad M 2009 Defining comborbidity: implications for understanding health and health services (Annals of Family Medicine vol. 7) pp 357-363

[13] Rockall T A Logan R F A Devlin H B Northfield T C and the steering committee and members of the National Audit of Acute Upper Gastrointestinal Haemorrhage 1996 Risk assessment after acute upper gastrointestinal haemorrhage (Gut vol 38) pp 316-321

[14] Jairath V Rehal S Logan R Kahan B Hearnshaw S Stanworth S and Travis S 1993 Acute variceal hemorrhage in the United Kingdom: Patient characteristics, management and outcomes in a nationwide audit (Digestive and Liver Disease vol. 46 pp 419–426.