IOSR Journal of Dental and Medical Sciences (IOSR-JDMS)
e-ISSN: 2279-0853, p-ISSN: 2279-0861.Volume 17, Issue 11 Ver. 8 (November. 2018), PP 68-78
www.iosrjournals.org

Model Selection and Model Averaging on Mortality of Upper Gastrointestinal Bleed Patients

Khuneswari Gopal Pillay[1*], SitiAisyahMohd Padzil[1],
RohayuMohd Salleh[1], Noraini Abdullah[2]

[1](*Department of Science and Mathematics, Faculty of Applied Science and Technology, UniversitiTun Hussein Onn Malaysia, PagohKampus, KM 1 Panchor Road, 84000, Muar, Johor, Malaysia)*
[2](*Mathematics with Economics Programme, Faculty of Science and Natural Resources, Universiti Malaysia Sabah, Jln. UMS, 88400 Kota Kinabalu, Sabah, Malaysia)*
Corresponding Author: Khuneswari Gopal Pillay

Abstract: *Model Selection (MS) is known to produce uncertainty into model-building process. Besides that, the process of MS is complex and time consuming. Therefore, Model Averaging (MA) had been proposed as an alternative to overcome the issues. This research will provide guidelines of obtaining best model by using two modelling approaches which are Model Selection (MS) and Model Averaging (MA), and then compares the performance of both methods. Corrected Akaike Information Criteria (AIC_c) and Bayesian Information Criteria (BIC) were applied in the model-building using MS to help determine the best model. In MA process, model selection criteria are needed to compute the weights of each possible model. Two model selection criteria ($AIC_c$and BIC) were compared to observe which approach will produce a model with a better performance. For guidelines illustration, data of Upper Gastrointestinal Bleed (UGIB) were explored to identify influential factors which led to the mortality of patients. At the end of the study, best model using MA was shown to have a better performance and AIC_c was proven to be a better model selection criterion approach in MA. In conclusion, the most significant factors for mortality of UGIB patients were identified to be shock score, comorbidity and rebleed.*

Keywords:*mortality, model-building, multiple binary logit, model selection criteria, model averaging*

Date of Submission: 14-11-2018 Date of acceptance: 29-11-2018

I. Introduction

Model Selection (MS) and Model Averaging (MA) is a method to produce practical models in applied research. The process of obtaining best model using MS requires elimination of insignificant variables. The goal is to omit insignificant variables one by one until only significant variables are left in the final model. Model selection criteria use maximum likelihood scores as a measure of fit. It is used to rank multiple competing models in term of information loss [1]. In this study, the model selection criteria used are corrected Akaike Information Criteria (AIC_c) and Bayesian Information Criteria (BIC). Model selection is well known for introducing additional uncertainty into the model building process. Hence, MA had been proposed as an alternative to model selection [2].

MA average the weight of each possible model using model selection criteria, to obtain the coefficient estimates on the weaker term so that the best model will yield a better prediction. Previously, [3]had suggested AIC to compute the weights of all possible models. Since there has been an issue when using AIC on small sample sizes as it will lead to a high degree of negative bias, AIC_cwhich was proposed by [4]is thus used instead. BIC which is proposed by [5] will also be applied in the MA process so as to compare the performances of both criteria.

This research will provide guidelines for model-building of Multiple Binary Logit (MBL) model as well as will compare the performance of best model for both approaches of Model Selection and Model Averaging. For MA, the performance of AIC_c and BIC in obtaining the weights to build the best model will also be compared. In order to illustrate the application of the proposed guidelines, data of Upper Gastrointestinal Bleed (UGIB) patients retrieved from Queen Elizabeth Hospital in Sabah were used. Based on the final best model, the factors influencing the survivability of UGIB patients were highlighted and discussed.

II. Methodology

2.1 Multiple Binary Logit

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). According [6], there are two main objectives of MBL which are prediction and providing useful information based on relationship between variables. The general MBL model as suggested by [7]is:-

$$Y_i = \beta_0 + \beta_1 X_{1i} + \beta_2 X_{2i} + ... + \beta_q X_{qi} + \mu_i \quad (1)$$

and

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

where the binary dependent variable is denoted by $Y_i$where $i = 1, 2, ..., n$,$X_j$is theX^{th}independent variable where $j = 1, 2, ..., q$, with the constant term of the model is denoted byβ_o, β_j is thej^{th} coefficient of independent variable where $j = 1, 2, ..., q$, μ is the random error of the model, and P_i is the probability of event occurs.Since the dependent variable is binary, the predicted Y will be in probability. As an example, a probability of 0.80 means that there is 80% chance of outcome 1 (success) to occur, and vice versa. The predicted probability can be computed using Equation (3) below.

$$E[Y_i] = P_i = \frac{e^{\beta_0 + \beta_1 X_{1i} + \beta_2 X_{2i} + ... + \beta_q X_{qi}}}{1 + e^{\beta_0 + \beta_1 X_{1i} + \beta_2 X_{2i} + ... + \beta_q X_{qi}}} \quad (3)$$

2.2 Guidelines on Model Selection

[8] presents the main phases of model-building for MBL model using MS as in Figure 1.

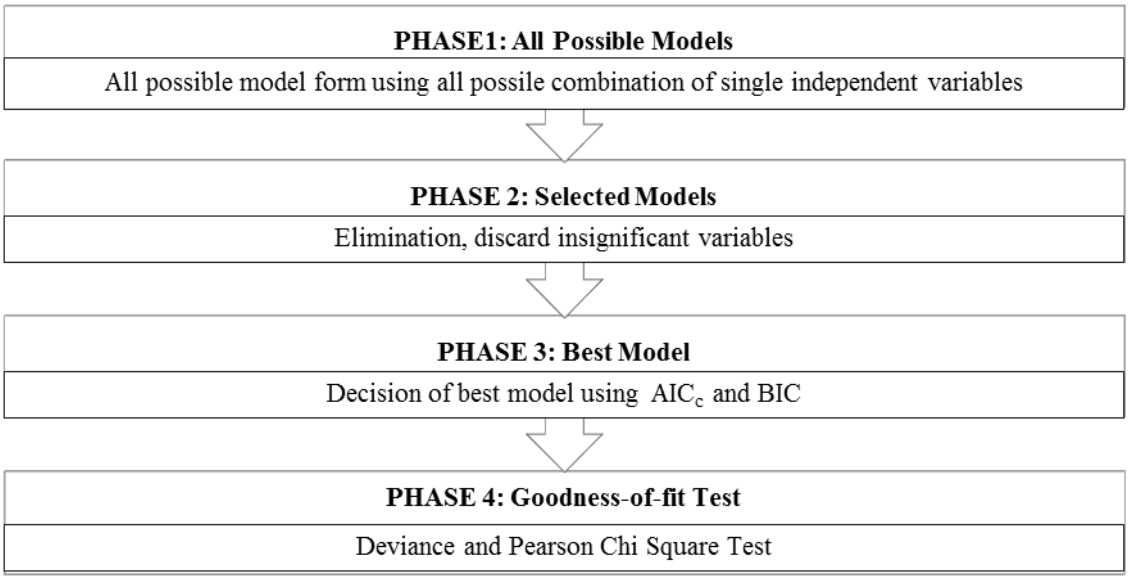

Fig. 1. Four Phases of Model-Building Approach using MBL

The first step to build a model is to list out all possible models form from all possible combinations of covariates. The formula to obtain the total number of all possible models without including interaction variable is as in Equation (4).

$$N = \sum_{j=1}^{q} (({}^{q}C_j)) \quad (4)$$

where N is the total number of all possible models, and q is the number of single independent variables, and $j = 1, 2, ..., q$.

After listing all possible models, the next step is to omit models with insignificant variables. The aim in *Phase 2* is to obtain the list of selected models which are models consist of only significant variables. The coefficient for all possible models were estimated using Maximum Likelihood (ML). For elimination of insignificant variables two test were conducted which are:

i. Coefficient Test
Eliminate any insignificant (non-contributing) variables

ii. Wald Test
Carried out to justify the removal of insignificant variables.

Alternatively, the elimination of insignificant variable can be done by comparing the p-value for each variable in the model. Variable with p-value more than 0.05 indicate an insignificant variable. The elimination 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 and the processes of omitting variables continue until all insignificant variables are eliminated. This process is conducted for each possible model. Normally, more than half of all possible models will be eliminated in this process, and the remaining models are then listed in selected models of *Phase 2*.

In *Phase 3*, the selected models are rank using model selection criteria and log-likelihood to determine the best model. Model with the least value of model selection criteria and maximum value for log-likelihood would indicate as the best model. Two model selection criteria are used in this study which are $AIC_c$and BIC.

$$AIC_c = -2\log L(M) - 2(q+1)\frac{n}{n-k-2} \quad (5)$$

$$BIC = -2\log L(M) - (q+1)[\log(n)] \quad (6)$$

where *L(M)* is minimum value for likelihood function of model *M*, *N* is number of observations and (q+1) is the number of parameters.

Finally, *Phase 4* in model-building is conducted to ensure the model validity and appropriateness. This research applies two goodness-of-fit tests, as suggested by [7] which are Pearson Chi-Square goodness-of-fit test and Deviance goodness-of-fit test. As supporting evidences, three scatter plots of residuals are plotted which are ordinary residuals against estimated probability, Pearson residuals against estimated probability and Deviance residuals against estimated probability. Residuals scatter plot for best model should approximately result in horizontal line with zero intercept.

i. Pearson Chi-Square Test
The hypothesis testing on 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 formula for test statistic and critical value for Pearson Chi-Square Test are in Equation (7) and Equation (8) respectively:

$$X_{P_i}{}^2 = \sum_{i=1}^{n} \frac{(Y_i - P_i)^2}{\sqrt{P_i(1-P_i)}} \quad (7)$$

$$\chi^2_{\text{critical}} = \chi^2_{(1-\alpha n-q-1)} \quad (8)$$

where $(Y_i - \hat{P}_i)$ are the ordinary residuals, $(\sqrt{P_i(1-\hat{P}_i)})$ is the estimated standard error of Y_i for $i = 1, 2, ..., n$, and $\hat{P}_i$ is the estimated probability.

ii. Deviance Test
The hypothesis testing on 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 formula for test statistic and critical value for Deviance Test are in Equation (9) and Equation (10) respectively:

$$G^2 = -2\sum_{i=1}^{n}[Y_i(\ln\hat{P}_i) + (1-Y_i)\ln(1-\hat{P}_i] \quad (9)$$

$$\chi^2_{\text{critical}} = \chi^2_{(1-\alpha n-q-1)} \quad (10)$$

To obtain the critical value, α is set as 5% or 0.05. If the value for teststatistic $\leq \chi^2_{\text{critical}}$, then accept the null hypothesis. The model is valid or appropriate when the null hypothesis is accepted.

2.3 Guidelines on Model Averaging

Figure 2 summarizes the five procedures of obtaining the best model of MBL using MA. *Phase 1* in MA is similar as in MS where the aim is to list all possible models that can be formed from the combination of independent variables. The computation for the total number of all possible models is as in Equation (4).

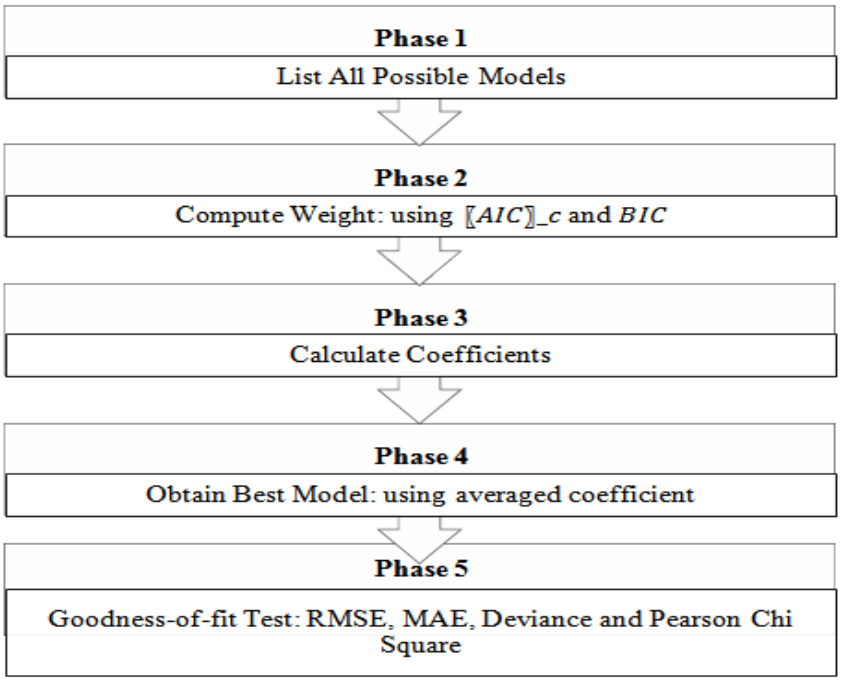

Fig 2. Five Phases of Model-Building using MA

In *Phase 2*, weights are calculated for each possible model. This research are comparing the performance of $AIC_c$and BIC to decide which criterion works the best in MA. Therefore, for each possible model, weights are calculated using both model selection criteria. The formula to obtain the weight is in Equation (11).

$$W_m = \frac{\exp(\frac{I_m}{2})}{\sum_{m=1}^{M} \exp(\frac{I_m}{2})} \tag{11}$$

W_m represents the weight for possible models, where $m = 1, 2, 3, ..., M$, and I_m is the model selection criteria as in Equation (5) and Equation (6). Table 1 presents the weights for all possible models.

Table 1. Weights for all possible models

MODEL	AIC_c	BIC	W_{AIC_c}	W_{BIC}
M1	AIC_{c_1}	BIC_1	$W_{AIC_{c_1}}$	W_{BIC_1}
M2	AIC_{c_2}	BIC_2	$W_{AIC_{c_2}}$	W_{BIC_2}
M3	AIC_{c_3}	BIC_3	$W_{AIC_{c_3}}$	W_{BIC_3}
. . .	. . .	. . .	. . ..	. . .
M_m	AIC_{c_M}	BIC_M	$W_{AIC_{c_M}}$	W_{BIC_M}

TOTAL			$\sum_{m=1}^{M} W_{AIC_c} = 1$	$\sum_{m=1}^{M} W_{BIC} = 1$

These weights will be used to compute the coefficients for the best model in *Phase 3*. The formula is shown in Equation (12):

$$\hat{\beta}_q = \sum_{m=1}^{M} W_m \hat{\beta}_{(q,M)} \tag{12}$$

, where $\hat{\beta}_{(p,M)}$ is the estimate of β_p under model Mfor $m = 1,2,3,...,M$. For an example, to obtain the estimated β_0, the calculation is as follows:

$$\hat{\beta}_0 = \beta_{(0,1)} W_1 + \beta_{(0,2)} W_2 + \beta_{(0,3)} W_3 + ... + \beta_{(0,M)} W_M$$

Each estimated coefficient $\ddot{\beta}_p$ will be calculated based on the coefficient value from each possible model as well as the weight obtained from each model.

In *Phase 4*, the best model is obtained after all the coefficients $(\beta_0, \ddot{\beta}_1, \ddot{\beta}_2, \ddot{\beta}_3, ..., \ddot{\beta}_q)$ are calculated. In order to choose the best model form using AIC_c and BIC, the Root Mean Square Error (RMSE) and the Mean Absolute Error (MAE) are computed. The model with minimum values for both accuracy measures would be selected as the best model. The formula to obtain RMSE and MAE as suggested by [9] are in Equation (13) and (14) respectively.

$$RMSE = \sqrt{\frac{\sum_{i=1}^{N} (\hat{Y}_i - Y_i)^2}{N}} \tag{13}$$

$$MAE = \frac{\sum_{i=1}^{N} \left| \hat{Y}_i - Y_i \right|}{N} \tag{14}$$

where, *N* is the total number of sample, Y_i is the actual value of dependent variables and $\ddot{Y}_i$ is the estimated value of *Y*.

Finally, in *Phase 5*, goodness-of-fit test is conducted to ensure the validity and appropriateness of best model using MA. This procedure is the same as goodness-of-fit test in MS where Pearson Chi-Square goodness-of-fit test and Deviance goodness-of-fit test are computed as well as three scatter plots of residuals are plotted.

2.4 Model Selection versus Model Averaging

In order to decide which modelling approach will produce a better performance model, the value for RMSE and MAE for both best models are computed and compared. The decision is to select the model with the minimum value of both accuracy measures. RMSE and MAE are computed using Equation (13) and Equation (14) respectively.

III. Data: Upper Gastrointestinal Bleed

Upper Gastrointestinal Bleeding (UGIB) is a medical term for bleeding that occurs in the esophagus, stomach or duodenum. Previously, [10] had introduced a scoring system to measure the risk of mortality for UGIB patients. Table 2 shows the Rockall scoring system which had included six variables. Rockall Score is a total score computed based on the score of each variable that were determined by the patient's condition. The greater the Rockall score, the higher is the risk of mortality.

Table 2: Rockall Scoring System

Variable	**Score**			
	0	**1**	**2**	**3**
Age	<60	60-79	>80	-
Shock	No shock	Tachycardia	Hypotension	
Comorbidity	No major comorbidity	-	Cardiac Failure, IHD, any major comorbidity	Renal failure, liver failure, disseminated malignancy

Diagnosis	Mallory-Weiss tear, no lesion	All other diagnosis	Malignancy of UGIT	-
Major Sign	None or dark spot	-	Blood in UGI tract, visible or spurting vessel	-

In this research, data of UGIB patients retrieved from the Queen Elizabeth Hospital in Sabah, Malaysia were studied to illustrate the procedures of obtaining the best MBL model. The data consisted of 410 samples. Seven covariates were chosen based on Rockall scoring system which were age, shock, comorbidity, diagnosis, major sign as well as Rockall score, and rebleed were analyzed to determine which factor contributed the most to predict mortality of UGIB patients. Table 3 presents the seven covariates in the study.

Table 3. Variable Descriptions

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 UGIT
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

IV. Data Analysis

4.1 Model-building using Model Selection

There were seven independent variables with a dependent variable been studied to obtain the best model using MS. In the first phase, by using Equation (4) the total number of all possible models when seven independent variables are included is 127 models, as shown in Table 4 below.

Table 4. 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	

In *Phase 2*, 90% of the listed possible models had been removed. The criteria for a model to be removed is possessing at least one insignificant variable, indicated by a p-value of more than 0.05. Each possible model was checked one by one, and the process of eliminating insignificant variable only allowed one independent variable to be removed at a time even though the model possesses more than one insignificant variable. Table 5 illustrated the process of insignificant variables elimination for model M120 which consisted of six independent variables:

$$\text{M120:} \quad Y = \hat{\beta}_0 + \hat{\beta}_1 X_1 + \hat{\beta}_2 X_2 + \hat{\beta}_3 X_3 + \hat{\beta}_4 X_4 + \hat{\beta}_5 X_5 + \hat{\beta}_6 X_6$$

Table 5. Elimination process in MS

Variables in model M120	P-value			
	M120.1	M120.2	M120.3	M120.4
X_1	0.5309	**0.5476**	-	-
X_2	0.0678	0.0650	0.0680	0.0427
X_3	0.0013	0.0014	0.0016	0.0012
X_4	**0.7056**	-	-	-
X_5	0.2534	0.2548	**0.2601**	-
X_6	0.0108	0.0084	0.0085	0.0031

The elimination in M120.1 starts from variable with highest number of p-value which is more than 0.05 (highlighted: 0.7056). Then the model is rerun and the same process is repeated until no more insignificant variable left in the model. Table 6 shows the remaining model (consist only significant variables) from 127 possible models. It can be seen that Model M120.4 with remaining variables (X_2, X_3 and X_6) is the same as model M46 as listed in all possible models.

Table 6. AIC_c, BIC and Log-Likelihood for selected models

Model	Selected Models	AIC_C	BIC	Log-Likelihood
M2	$\hat{Y}_2 = \hat{\beta}_0 + \hat{\beta}_2 X_2$	-297.720	- 285.731	151.8895
M3	$\hat{Y}_3 = \hat{\beta}_0 + \hat{\beta}_3 X_3$	-305.9812	-293.9919	156.0202
M5	$\hat{Y}_5 = \hat{\beta}_0 + \hat{\beta}_5 X_5$	-295.2531	-283.2638	150.6561
M6	$\hat{Y}_6 = \hat{\beta}_0 + \hat{\beta}_6 X_6$	-307.0646	-291.0987	157.5817
M7	$\hat{Y}_7 = \hat{\beta}_0 + \hat{\beta}_7 X_7$	-316.6008	**-300.6349**	162.3498
M14	$\hat{Y}_{14} = \hat{\beta}_0 + \hat{\beta}_6 X_6 + \hat{\beta}_7 X_7$	-311.6363	-295.6704	159.8675
M17	$\hat{Y}_{17} = \hat{\beta}_0 + \hat{\beta}_2 X_2 + \hat{\beta}_3 X_3$	-302.1722	-290.1828	154.1157
M21	$\hat{Y}_{21} = \hat{\beta}_0 + \hat{\beta}_2 X_2 + \hat{\beta}_6 X_6$	-304.3846	-292.3952	155.2219
M28	$\hat{Y}_{28} = \hat{\beta}_0 + \hat{\beta}_3 X_3 + \hat{\beta}_6 X_6$	-309.6446	-293.6787	158.8717
M46	$\hat{Y}_{46} = \hat{\beta}_0 + \hat{\beta}_2 X_2 + \hat{\beta}_3 X_3 + \hat{\beta}_6 X_6$	**-319.8614**	-299.9291	**165.0049**
M59	$\hat{Y}_{59} = \hat{\beta}_0 + \hat{\beta}_3 X_3 + \hat{\beta}_6 X_6 + \hat{\beta}_7 X_7$	-310.2039	-290.2716	160.1762

Two model selection criteria which are AIC_c and BIC, and log-likelihood values are observed in *Phase 3* to help in choosing the best model. From Table 6 above, model M46 has shown to have the minimum value of AIC_c (-319.8614), and maximum value of log-likelihood (165.0049). Even though model M7 has the minimum value for BIC (-300.6349), the difference was intangible. Therefore, model M46 is concluded as the best MBL model for UGIB patients, given by:-

$$\text{M46:} \quad \hat{Y}_{46} = \hat{\beta}_0 + \hat{\beta}_2 X_2 + \hat{\beta}_3 X_3 + \hat{\beta}_6 X_6$$

To ensure the validity of model M46, goodness-of-fit test is conducted. The hypothesis testing for both the Pearson Chi-Square and Deviance tests are as follows:

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

For Pearson Chi-Square test, the value of test statistics and critical value are $\chi_{r_{P_i}}^2 = 258.3356$ and $\chi^2_{(0.95,406)} = 360.293$. Since 258.3356<360.293, H_0 is accepted. Therefore, it can beconcluded that best model M46 is appropriate and valid.

For Deviance test the test statistic G^2 is 86.73909, and for $\chi^2_{critical}$, the value is $\chi^2_{(0.95,406)} = 348.347$. Since 86.73909 <348.347, H_0 is accepted. Therefore, it can beconcluded that best model M46 is appropriate. Figure 3, 4 and 5 show the scatter plots of residuals.

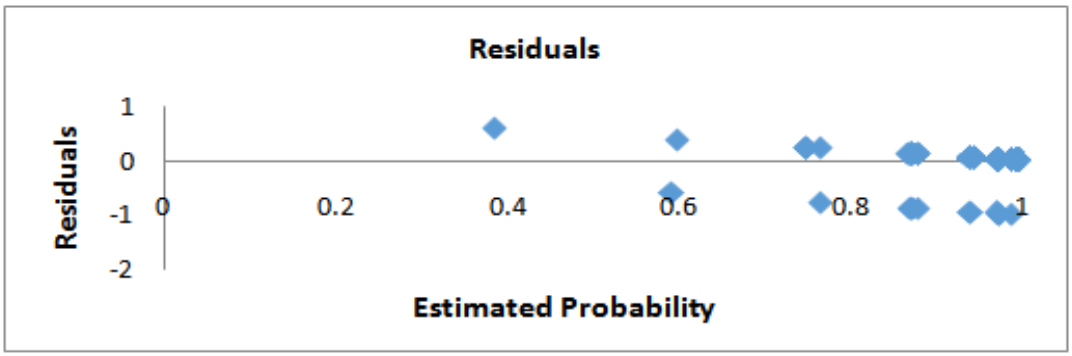

Figure 3. Ordinary Residuals for model M46.

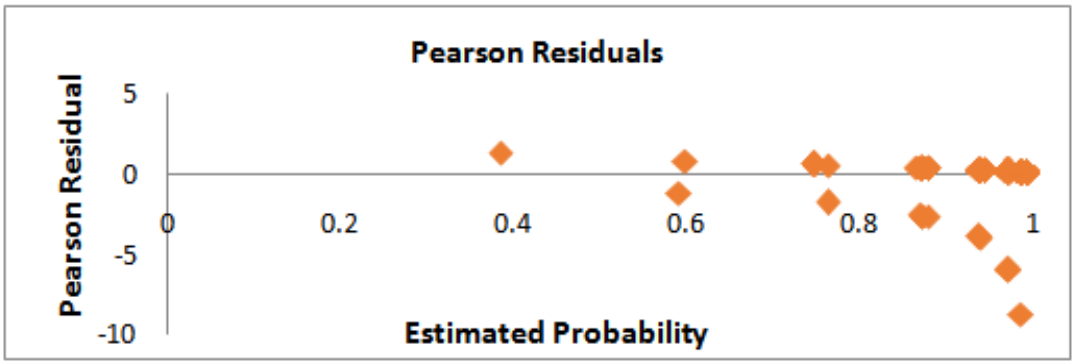

Figure 4. Pearson Residuals for model M46.

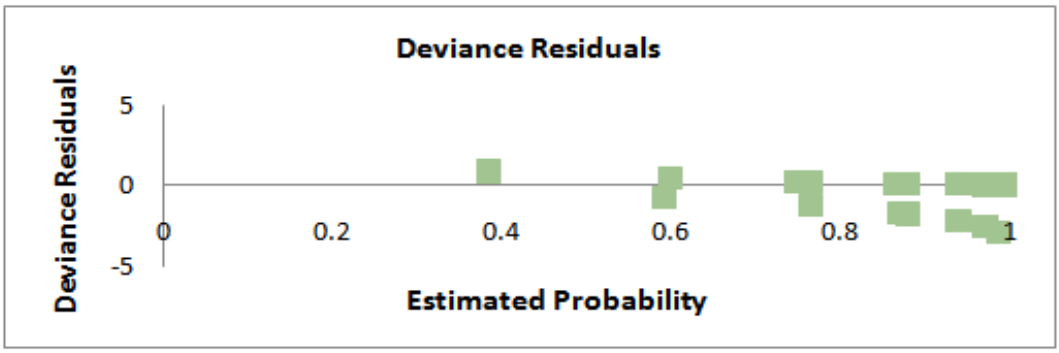

Figure 5. Deviance Residuals for model M46.

The ordinary residuals plot and Deviance residuals approximately result in horizontal line and zero intercept which justify the validity of the best model. Even though Pearson residuals do not follow the horizontal line but the plots do have zero intercept. According to [11], 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 UGIB patient's data as [7] had 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 UGIB patients. The best model of UGIB patients using MS is

$$\ddot{Y}_{46} = 0.464 - 0.837X_2 - 0.808X_3 + 2.361X_6$$

Table 7 shows the results of coefficient and p-value for model M46.

Table 7. Model M46 results

Variable	Coefficient	p-value
Intercept	0.464	$2x10^{-16}$
X_2	-0.837	0.0427
X_3	-0.808	0.0012
X_6	2.361	0.0031

4.2 Model-building using Model Averaging

First phase of MA is similar as in MS where 127 possible models are listed. Then the weight for each possible models are calculated using Equation (13) where the$I_m$is AIC_c and BIC. Table 8 shows the best model of MA using AIC_c and BIC.

Table 8. Best Model of MA using AIC_c and BIC

Model	I_m	Model	RMSE	MAE
MA1	AIC_c	$\ddot{Y} = 0.7652 + 0.0089X_1 - 0.0271X_2 - 0.0288X_3 - 0.0067X_4 - 0.0073X_5 + 0.1299X_6 - 0.0263X_7$	0.1642	0.0798
MA2	BIC	$\ddot{Y} = 0.7580 + 0.0074X_1 - 0.0289X_2 - 0.0307X_3 - 0.0085X_4 - 0.0096X_5 + 0.1332X_6 - 0.0354X_7$	0.1699	0.0982

Since the both model selection criteria (AIC_c and BIC) are studied in modelling using MA, two best model are obtained (one from each criteria). RMSE and MAE are computed to observe which model selection criteria would produce a better performance model in MA. From Table 8, it can be seen that the best model usingAIC_cshows a better performance as it has a lower error value. Therefore, model MA1 is chosen as the best model.

Goodness-of-fit test is conducted using the final best model. The hypothesis testing for both Pearson Chi-Square and Deviance goodness-of-fit test is as follows:

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

For Pearson Chi-Square test, the value of test statistics and critical value are $\chi_{r_{P_i}}{}^2 = 229.8983$ and $\chi^2_{(0.95,402)} = 356.254$ respectively. Since 229.8983<356.254, H_0 is accepted. For Deviance test, the test statistic G^2 is 290.6504, and for $\chi^2_{critical}$ the value is $\chi^2_{(0.95,402)} = 356.254$. Since the test statistic< critical value, H_0 is accepted. Therefore, it can be concluded that best model is appropriate.

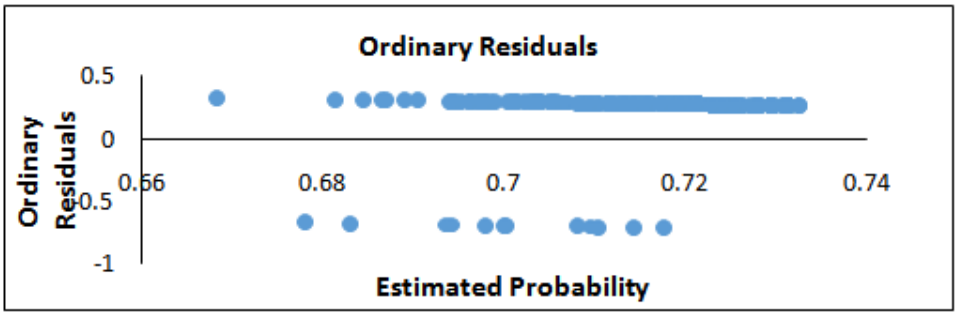

Figure 6. Ordinary Residuals.

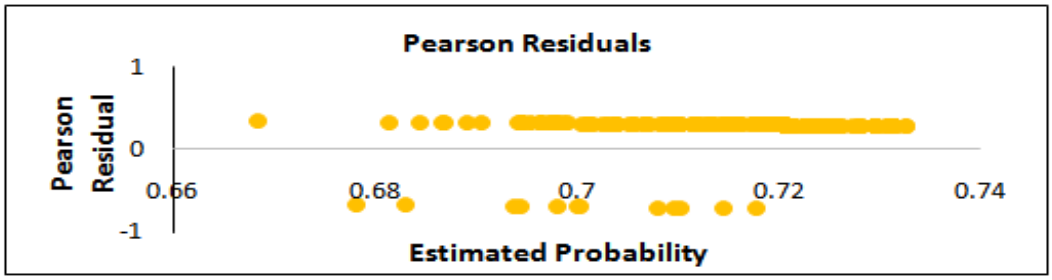

Figure 7. Pearson Residuals.

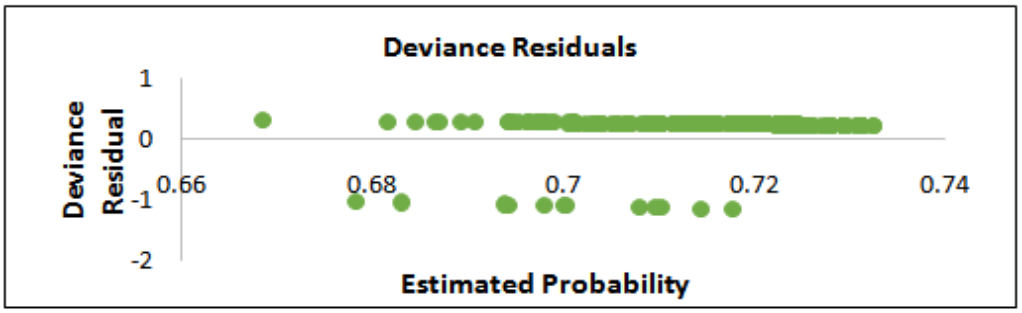

Figure 8. Deviance Residuals.

Figures 6, 7 and 8respectively present the three scatter plot of residuals (Ordinary residuals, Pearson Residuals, and Deviance Residuals). All three scatter plots show horizontal lines with zero intercept which supported the goodness-of-fit results where the final model is valid and appropriate. As a conclusion, the fitted best model of UGIB patients formed using MA is given as:-

$$\ddot{Y} = 0.7652 + 0.0089X_1 - 0.0271X_2 - 0.0288X_3 - 0.0067X_4 - 0.0073X_5 + 0.1299X_6 - 0.0263X_7$$

4.3 Comparison between Model Selection and Model Averaging

Table 9 shows the comparison of accuracy measure between best model form using MS and MA.

Table 9. RMSE and MAE for best models using MS and MA

Methods	Models	RMSE	MAE
Best model using MS	$\ddot{Y} = 0.464 - 0.837X_2 - 0.808X_3 + 2.361X_6$	3.1821	3.315
Best model using MA	$\ddot{Y} = 0.7652 + 0.0089X_1 - 0.0271X_2 - 0.0288X_3 - 0.0067X_4 - 0.0073X_5 + 0.1299X_6 - 0.0263X_7$	**0.1642**	**0.0789**

From the results above, modelling using MA has shown to produce a model with a better performance compared to MS. Therefore, the best model of UGIB patients is:-

$$\ddot{Y} = 0.7652 + 0.0089\text{AgeScore} - 0.0271\text{ShockScore} - 0.0288\text{Comorbidit} - 0.0067\text{Diag.Score} - 0.0073\text{MajorSign} + 0.1299\,\text{Re bleed} - 0.0263\text{RockallGroup}$$

Table 10 shows the results of coefficient and p-value of the final fitted best model.

Table 10. Coefficients and p-values of the best model

Variable	Coefficient	p-value
Constant	0.7652	2e-16
X_1	0.0089	0.5636
X_2	**-0.0271**	**0.0439**
X_3	**-0.0288**	**0.0012**
X_4	-0.0067	0.7411
X_5	-0.0073	0.5016
X_6	**0.1299**	**0.0009**
X_7	-0.0263	0.2423

From the best model, when the values for all the variables in the model are 0, the probability of UGIB patient's survivability is $P_i = \frac{\exp^{0.7652}}{1+\exp^{0.7652}} = 0.6825 \approx 0.68$

If there is one-unit increase in shock score (X_2), the $\hat{Y}$ will decrease by 0.0271. Therefore, the probability of UGIB patient's survivability will decrease. Similarly, the probability of UGIB patient's survivability also will decrease by 0.0288 if there is one-unit increase in comorbidity (X_3). The probability of UGIB patient's survivability will decrease by 0.1299 if there is one-unit increase in rebleed (X_6). Whereas, the probability of UGIB patient's survivability will increase by $2 \times 0.1299 = 0.2598$, if there is no rebleed.

V. Discussions and Conclusion

The guidelines for MS and MA methods proposed in this research can be used on any data with binary dependent variable. From the analysis, MA methods is proven to produce model with better performance. The results of accuracy measure show a huge difference between best model obtain using MS and MA. The RMSE and MAE value for best model using MS is shown to be almost twenty times larger when compared with best model of MA.

In MA methods, two model selection criteria are applied in *Phase 2* to compute weights for each possible models. Study by [3] suggested the use of AIC to compute weight in MA, but in this research AIC_c and BIC are tested and the performance are measured using RMSE and MAE. It is a good idea to compare several model selection criteria with different performance measures so as to be more accurate in making the final decision to choose the best model. From the analysis, AIC_c has proven to yield a model with lower RMSE and MAE which indicate a better performance when compared with BIC.

Based on the final best model, only three variables are appeared to be significant in the best model where the p-value is less than 0.05. Therefore, it can be concluded that the factors affecting the survivability of UGIB patients are found out to be shock score (X_2), comorbidity (X_3) and rebleed (X_6).

References

[1]. M.R.E.Symonds, and A.Mousalli. (2011). A Brief Guide to Model Selection, Multimodel Inference and Model Averaging in Behavioural Ecology Using Akaike's Information Criteria. *Behavioural Ecology Sociobiology*, 65, pp.13-21.

[2]. G.Claeskens, and N.L.Hjort. (2008). *Model Selection and Model Averaging*. United Kingdom: University Press, Cambridge.

[3]. T. S.Buckland, K.P.Burnham, andN.H.Austin. (1997). Model Selection: An integral partof inference. *Biometrics*, 53(2), pp. 603-618.

[4]. C.M. Hurvich, and C.L. Tsai, Bias of the corrected AIC criterion for underfitted regression and time series models. (1991).*Biometrika* 78(3),499-509.

[5]. G.Schwarz. Estimating the dimension of a model. (1978). *Annals of Statistics* 6, 461–464.

[6]. D.W. Hosmer, and S.Lemeshow, (2000). Applied Logistic Regression. 2nd edition. United States: John Wiley and Sons Inc.

[7]. M.H.Kutner, C.J.Nachtsheim, and J.Neter.Applied Linear Regression Models. 4th edition.(McGraw-Hill Inc.: Singapore, 2008).

[8]. S.M.P.Aisyah, and G.Khuneswari. (2017). *Model Building Approach using Multiple BinaryLogit Model for Categorical and Continuous Variables.*UniversitiTun Hussein Onn Malaysia. Master's Thesis.

[9]. T.Chai, and R.R.Drexler. (2014). Root Mean Square Error or Mean Absolute Error?Arguments against avoiding RMSE in the literatures. *Geoscientific Model Development*, 7, pp.1247-1250.
[10]. T.A.Rockall, R.F.A.Logan, H.B.Devlin, T.C.Northfield, and the steering committee and members of the National Audit of Acute Upper Gastrointestinal Haemorrhage. (1996). Risk assessment after acute uppegastrointestinalhaemorrhage. *Gut*, 38, 316-321.
[11]. A.Noraini, H.J.Zainodin, and L.B.Rick. (2013). Risk factor determination on UGIB patients in Kota Kinabalu, Sabah Malaysia. *Journal of Medical Sciences*, 13(7), 526-536. Doi:10.3923/jms.2013.

Khuneswari Gopal Pillay, "Model Selection and Model Averaging on Mortality of Upper Gastrointestinal Bleed Patients"IOSR Journal of Dental and Medical Sciences (IOSR-JDMS), vol. 17, no. 11, 2018, pp 68-78.