Sindh Univ. Res. Jour. (Sci. Ser.) Vol.44 (3) 511-516 (2012)

SINDH UNIVERSITY RESEARCH JOURNAL (SCIENCE SERIES)

Statistical Models For Predicting Intensive Care Unit Mortality; Comparison And Choice

I. A. ARSHAD, F. RASHEED* A.W. SHAIKH**

Department of Mathematics and Statistics, International Islamic University, Islamabad,

Received 2nd April 2012 and Revised 11th June 2012

Abstract: The most important clinical outcome measure from the patient's perspective is ICU mortality. ICU mortality provides a global impression of the ICU performance but is affected by a large number of factors such as severity of illness and age of patients. Special efforts can be made by using ICU prediction model to prevent the mortality risk factors and to improve critical care medicine in Pakistan. Mortality rate in medical intensive care unit was found 88.33% and a high mortality rate was observed for all age groups indicating poor socio-economic and health conditions. Five predictors unconsciousness, history of pre-existent chronic disease, creatinine, SPO2 and temperature were obtained by using both best subset logistic regression and stepwise logistic regression. The stepwise discriminant analysis is also applied and selected five predictors of logistic model and one additional predictor mechanical ventilation, although the assumption of normality and equal variances and covariance are violated. The correct classification rate (CCR) for logistic models and discriminant model were 90% and 80.7% respectively. The area under the receiver operating curve (ROC) for logistic models and discriminant model were 90% and 88.6% respectively. The logistic discrimination reported mechanical ventilation as insignificant factor and its CCR was equal to 90%, which was about 9.3% higher than discriminant model. The ROC curves for logistic model and logistic discrimination model cover the same area. The comparison of selected models indicated the superiority of logistic model over discriminant model. This study shows that logistic model has greater classification and discriminating power than discriminant model when most of the predictors are attribute type. Thus when the assumption of multivariate normality and equal variance and co-variance are not met, and the purpose is classification then discriminant analysis can provide accurate results in terms of variable selected or the signs of the coefficients, but final model must be based on maximum likelihood estimators of the coefficients.

Keywords: Intensive Care Unit, Logistic Regression, Discriminant Analysis, Logistic Discrimination, Correct Classification Rate, Misclassification Rate, Receiver Operating Characteristic Curve.

1. INTRODUCTION

An intensive care unit, or ICU, is a specialized section of a hospital that provides comprehensive and continuous care for persons who are critically ill. Most hospitals operate an ICU to provide care for their most critically ill patients. But unlike other units of a hospital, the ICU is not devoted to a single disease but treats a variety of illnesses, all of which are life threatening diseases.

In this study, we are studying patients admitted in Medical ICU, which includes patients of coronary disease. ICU outcome can be seen as one or more events in the course of a disease process, such as morbidity, mortality, time to recovery from disease, or quality of life. Until now, in medical research, outcome prediction models Acute Physiology and Chronic Health Evaluation (APACHE), Simplified Acute Physiology Score (SAPS) and Mortality Probability Models (MPM) in intensive care focus exclusively on hospital mortality. MPM was developed after APACHE and SAPS to outcome prediction in ICU by using multiple logistic regression technique to choose and weigh the variables rather than by a consensus from a panel of expert. In this study, the purpose is to determine MPM. ICU mortality is defined as death within the stay of medical ICU. The patients who discharged alive and subsequently die are not considered as death. In this study ICU mortality is considered as a binary variable. Information from ICU patients can be collected at four different periods. Which are, i) At the time of admission in ICU ii) After 24 hours iii) After 48 hours iv) After 72 hours.Risk factors for ICU mortality are different for these periods. In this study we are using only demographic and condition variables at the time of ICU admission for modeling process.

Logistic regression is used to determine relationship between a binary response variable and one or more predictor variables, which may be either discrete or continuous. Binary outcome data is common in medical applications. Chin (2001) has also reported a significant increase in the use of multivariable logistic regression in the public health and epidemiologic research literature. Kellie *et al.* (2007) reported the use and importance of logistic regression models in epidemiological studies for estimating associations that demographic, behavioral, and risk factor variables have on a dichotomous outcome, such as disease being present versus absent.

An alternative method to logistic Regression is Discriminant Analysis. Discriminant Analysis (DA) is a multivariate technique for investigating the relationship among several predictors and a single nominal response variable. It is used to determine which variables discriminate between two or more naturally occurring groups.

++ Corresponding author:I. A. ARSHAD *Email: irshad.ahmad@iiu.edu.pk*
*Department of Statistics, F.G. Postgraduate College for Women, Kashmir road, Rawalpindi.
**Institute of Mathematics and Computer Science, University of Sindh, Jamshoro.

Both methodologies have been extensively applied in research, especially in medical and sociological sciences. Like Logistic Regression, Discriminant analysis is also a classification method that is used to determine which set of variables discriminate between two or more naturally occurring groups and to classify an observation into these known groups. Thus, it can be proposed that both discriminant analysis and logistic regression can be used to predict the probability of a specified outcome using all or a subset of available variables.

Gavilan *et al.* (2007) applied the stepwise discriminant analysis to data retrospectively obtained from 243 patients who underwent surgery for cancer of the larynx. In this study, 62 variables were studied in each patient. The stepwise discriminant analysis 62 variables to 11 as important variables to obtain the result.

2. MATERIAL AND METHODS

About Data

This is a retrospective study of patients admitted to the ICU of department of medicine Unit-I and II ***Holy Family Hospital (HFH)*** Rawalpindi. The data of 300 patients consist of information on 26 variables, were collected from the records of HFH from Jan. 2007 to March 2008. The demographic and condition variables available at the time of ICU admission used in the modeling process were 1) Age 2) Sex 3) Socioeconomic Status 4) Location Before ICU 5) Cancer Part of Present Problem 6) History of Chronic Renal Failure 7) History of Pre-existent Chronic Disease 8) Infection Probable at ICU Admission 9) CPR Prior to ICU Admission 10) Systolic Blood Pressure at ICU Admission 11) Heart Rate at ICU Admission 12) Previous Admission to an ICU Within 6 Months 13) Temperature 14) History of Obstructive Pulmonary Disease 15) Accidental Cases 16) Acute Respiratory Failure 17) Multiorgan Failure 18) Need for Mechanical Ventilation 19) PO2 from Initial Blood Gases 20) PH from Initial Blood Gases 21) PCO2 from initial Blood Gases 22) SPO2 23) Bicarbonate from Initial Blood Gases 24) Creatinine 25) Level of Consciousness at ICU Admission. The response variable was the outcome of ICU patients. So response variable is the binary variable Y that takes 0 if the patient dies in the ICU and 1 otherwise.

3. METHODOLGY

Model Building Strategy

The main goal of traditional approach to statistical model building is to minimize the number of variables in the model and the resultant model must be numerically stable, and can easily be generalized. (Hosmer and Lemeshow 2000. In this study, model building strategy for LR proposed by Hosmer and Lemeshow is applied. Following are the steps involved in the selection of variables for ICU mortality prediction model. 1) Checking the assumptions of logistic regression model and discriminant model, which may affect the estimates and prediction, are used. 2) Univariate logistic regression is applied and the importance of each predictor is checked by Wald statistics. The predictors having p-value<0.25 are selected as candidate predictors. 3) Three variable selection procedures i) Best Subset Logistic Model ii)Stepwise Logistic Model iii) Stepwise Discriminant model are used. These three methods are then compared to determine best model for ICU mortality. 4) Checking possible interactions that have some clinical basis to interact with each other 5) Checking the importance of interaction terms added in the model. 5) Performing overall goodness-of-fit and model diagnostics for final ICU mortality Model.

The multiple Logistic regression was developed using BMDPLR (2009). This model has the following general form $P(y=1/x)=\pi(x)$. Where $\pi(x)$ has the form,

$$\pi(x)= \frac{\exp\{\beta_0 + \beta_1 x_1 + \beta_2 x_2 + \ldots + \beta_p x_p\}}{1+\exp\{\beta_0 + \beta_1 x_1 + \beta_2 x_2 + \ldots + \beta_p x_p\}}$$

To select the best subset logistic regression model, Mallows' C(q) criterion (an analog of Mallows Cp used in standard linear regression) proposed by Hosmer and Lemeshow (2000). The Mallow's Cq statistic for a subset of '*q*' of the '*p*' variables is

$$Cq = \frac{\chi^2 + \lambda^*}{\chi^2/(n-p-1)} + 2(q+1) - n$$

Where $\chi^2 = \sum \left\{ (Y_i - \hat{\pi}_i)^2 \Big/ [\hat{\pi}_i(1-\hat{\pi}_i)] \right\}$, the pearson chi-square statistic for the model with p variables and λ^* is the multivariable Wald statistic for the hypothesis that the co-efficients for the p-q variables not in the model are equal to zero. The second approach of variable selection used in this study is Stepwise Logistic Regression Model, which is a statistical algorithm that checks for "importance" of variables, and either includes or excludes them on the basis of a fixed decision rule. The third approach of variable selection is Discriminant Analysis, which is an alternative To Logistic Regression. The linear discriminant function (LDF) is represented by equation $LDF = b_0 + b_1 x_{i1} + b_2 x_{i2} + \cdots . + b_p x_{ip}$ .Where b_j is the value of the j^{th} coefficient, j = 1, …, p, and x_{ij} is the value of the i^{th} case of the j^{th} predictor. The stepwise discriminant analysis technique used to determine the weight and direction of each variable and how it contributes to group membership. The stepwise discriminant algorithm includes the predictor in each step, which minimize Wilk's Lambda. The Wilk's Lambda used in this

algorithm is a multivariate analysis of variance statistic that tests the equality of group means for the variables in the discriminant function. Wilk's Lambda follows the equation:

$$\Lambda = \left|\frac{S_{wg}}{S_{bg}+S_{wg}}\right|$$

Fisher's Linear discriminant function is optimal when the two populations have multivariate normal distribution with equal covariance matrices, but have been shown to be relatively roboust to departure from normality. In some cases, however, where the normality assumption is considered invalid it might be worth considering alternative methods: one which has been widely used is logistic discrimination. The essential feature of this approach is to assume the following form for the probabilities of class membership of two groups D_1 and

D_2. $P(D_1/x) = \frac{e^{\left(b_0+\sum_{j=1}^{p} b_j x_j\right)}}{1+e^{\left(b_0+\sum_{j=1}^{p} b_j x_j\right)}}$

The estimates $b_0, b_1, \ldots, b_p$ are estimated by the ML method. After estimation of the parameters, allocation of new individuals can be performed on the basis of scores given by $b_0 + \sum_{j=1}^{p} b_j x_j$. If this is positive the individual is allocated to D_1 (Since $P(D_1) > P(D_2)$), if negative to D_2

4. RESULTS AND DISCUSSIONS

ICU Results

Total 300 patients were submitted in medical ICU of HFH during the period Jan. 2007 to March 2008. Out of 300, 265 patients were died implied that 88.33% rate of mortality. The high rate of mortality indicates most of the ICU patients when admitted were critically ill. The different studies like Lemeshow *et al.* (1988), Nele *et al.* (1999) and Ashish *et al.* (2004) showed 19%, 22% and 33.7% ICU mortality rate respectively. These mortality rates are very low as compared to the mortality rate reported in the present study. Univariate logistic regression model is fitted at 0.25 level recommended by Mickey and Greenland (1989). Nine of the 25 predictors (SEX, HOSP, CAN, PRE, INF, HRA, PO2, PCO2, and PH) showed insignificant relationship with vital status.

Best Subset Logistic Regression

The best subset logistic regression technique is applied, by using only those predictors which are significant at univariate analysis. Commonly used statistical packages do not have program for directly obtaining the values of Mallow's C_q, to determine the best subset logistic model. Hosmer and Lemeshow (2001) proposed the use of linear regression program with pseudo-dependent variable $\mathbf{z_i}$ case weights $\mathbf{v_i}$ and predictors $\boldsymbol{x_i}$ to determine best subset logistic model. The general element of the pseudo-dependent variable is

$$z_i = \ln\left(\frac{\hat{\pi}}{1-\hat{\pi}}\right) + \frac{(Y_i-\hat{\pi})}{\hat{\pi}(1-\hat{\pi})}.$$

The values of z_i and the case weights $v_i = \hat{\pi}(1-\hat{\pi})$ are computed by the fitted values of the full model. The best subset selection program BMDP9R, is used to obtain the values of Mallow's Cq to identify best subsets of predictors for each q, q=1, 2,3,….,16. The results are shown in the following table.

Table 1- Five Best Subset Models identified Using Mallow's Cq

Model	Model Covariates	Cq	λ^*	df	P-value
1	CRE	0.72	70.69	15	<0.0001
2	CRE UNCONS	1.44	41.99	14	0.0001
3	PCD CRE UNCONS	1.89	32.42	13	0.0021
4	PCD SPO2 CRE UNCONS	2.19	25.10	12	0.0144
5	PCD TEMP SPO2 CRE UNCONS	3.00	19.26	11	0.0565

The above table showed, five best models are selected by the program BMDP9R. The p-values of λ^* for first four models are small enough to conclude that the predictors excluded from these models are significant and implied that some of the excluded predictors must be included in the model. Hosmer and Lemeshow (1989) suggested that good subsets of variables will be those with small values of Cq and Wald statistic for excluded predictors from the selected models should be insignificant. Finally they recommended that identification of a final best model is thus a combination of statistical and biological criteria. In table 1, CRE is in five, UNCONS in four, PCD in three and SPO2 appears in two of the five models. The p-value of λ^* only for model five is insignificant, implies none of the excluded predictor is significant. Therefore, model five is selected as best model statistically and clinically. This result showed that age is not an important predictor for ICU mortality.

Stepwise Logistic Regression Model at P_E=0.05 and
$-26.66 + 2.35UNCONS + 1.70PCD + 2.38CRE - 1.59SPO2 + 0.28TEMP$
$\hat{C}$ **(p-value)=** 5.3878(0.6127)
General R^2 (Max-Rescaled R^2)= 0.2107(0.4103)

Significance of age means, the risk of death increases with age, but in this study 77%, 93% and 95% mortality rate was observed in 91 young patients, 136 middle age patients and 73 old age patients respectively. This implied that rate of mortality in Pakistan is very high for all age groups. In HFH ICU data, among young age patients, 34(23 died) were the case of snake bite, poison and drug addiction(which immediately cause the problem of multiorgan failure), 10 (9 died) were the cases of COPD and 47 (37 died) were the cases of INF

like Tuberculosis, which are common risk factors of young age mortality in underdeveloped countries like Pakistan, so that's why our study showed different risk factors of ICU mortality from the risk factors of mortality showed by different researches conducted in developed countries.

Stepwise Logistic Regression

This algorithm was performed by BMDPLR program, which requires alpha level to enter and to remove a predictor. By default, BMDPLR used 0.10 and 0.15. But most of the medical studies used p-levels between 0.05 to 0.10. Glen *et al.* (1998) used 0.10 significance level to enter and to remove predictors in SLR. Fabian *et al.* (2005) used 0.05 significance level to enter predictors in forward stepwise LR. Levinson (2007) applied both SDA and SLR with 0.05 p-level to enter and 0.10 p- level to remove predictors. This study is related to medical and critical care research, where the main objective is to identify the risk factors of mortality or a specified disease, so the significance level between 0.05 to 0.10 for stepwise logistic regression is used. The following table showed the results of stepwise logistic regression.

The above table showed that five predictors are selected and all these predictors are significant at 5% level of significance. Hosmer and Lemeshow goodness of fit statistic $\hat{c}$ is insignificant indicates the adequacy of model.

Stepwise Discriminant Analysis

The stepwise discriminant Analysis was performed by using both BMDP7M program and SPSS 16.0. Both produced the same results but BMDP7M does not report Wilk's lambda at each step. These programs choose the variables used in computing linear discriminant function in a stepwise manner. The F-value to enter a predictor is 3.8416 (5%) and to remove is 2.7056(10%). The BMDP7M program and SPSS both substantially reduced the 25 possible admission condition and demographic predictors to 6 predictors. The 6 predictors are UNCONS, PCD, CRE, SPO2, MV and TEMP. Both programs enter the predictors in the same order. The LDF is given by

$$\boldsymbol{LDF} = -\mathbf{13.118} + \mathbf{1.736}\boldsymbol{UNCONS} + \mathbf{.986}\boldsymbol{PCD} + \mathbf{.929}\boldsymbol{CRE} - \mathbf{.813}\boldsymbol{SPO2} + \mathbf{.520}\boldsymbol{MV} + \mathbf{.113}\boldsymbol{TEMP}$$

The coefficients in LDF are the partial coefficients, reflecting the unique contribution of each predictor to the classification of the response variable. The predictor UNCONS, PCD, CRE, SPO2 and MV have the most contribution in the classification.

The classification rule is given as follows, which allocates an individual with UNCONS, PCD, CRE, SPO2, MV and TEMP to the mortality if

$$\boldsymbol{LDF} = -\mathbf{13.118} + \mathbf{1.736}\boldsymbol{UNCONS} + \mathbf{.986}\boldsymbol{PCD} + \mathbf{.929}\boldsymbol{CRE} - \mathbf{.813}\boldsymbol{SPO2} + \mathbf{.520}\boldsymbol{MV} + \mathbf{.113}\boldsymbol{TEMP} > \mathbf{0}$$

A plot of the discriminant scores for all individuals is shown in the following **(Fig.1).**

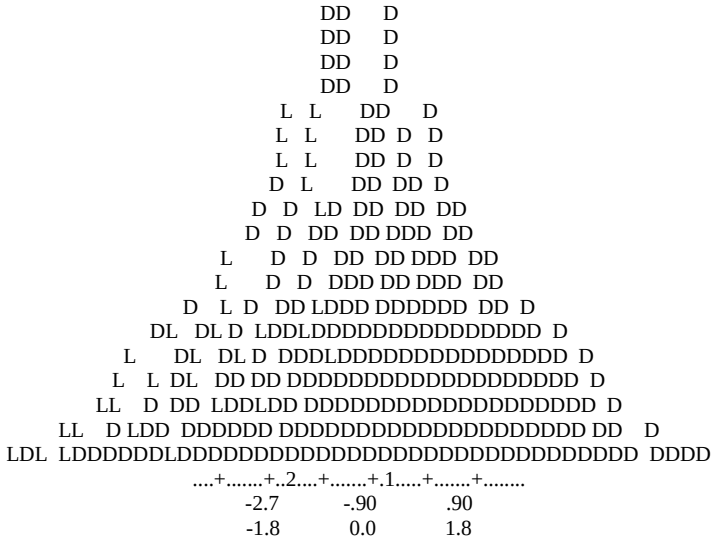


Fig1. Discriminant Score Plot For ICU Patients

The individual having discrminant score > zero assigned to group died and patients having discriminant score < zero assigned to group survived. In the above figure D denotes died and L denotes survived. The above plot showed a small overlap between two groups. Some died patients are incorrectly allocated to group survived.

Comparison Of Best Subset Logistic, Stepwise Logistic Regression And Stepwise Discriminate Models

The estimated coefficients of best subset logistic model, stepwise logistic model and stepwise discriminant analysis are presented in table 2 to compare these models on the following aspects; variables, number of variables, magnitude and signs of coefficients.

Table 2- Variables and Coefficient of Best Subset Logistic, Stepwise Logistic Regression and Stepwise Discriminant Analysis

Variables	Best Subset Logistic Regression	Stepwise Logistic Regression Coefficient	Stanadardized Discriminant Coefficient
UNCONS	.3495	2.3495	1.736
PCD	.6975	1.6975	0.986
CRE	.3826	2.3826	0.929
SPO2	1.5892	-1.5892	-0.813
TEMP	.2771	0.2771	0.113
MV			0.520
Intercept	26.6875	-26.6875	-13.0821

From the above table, both the best subset logistic model and stepwise logistic model produced the same predictors but six predictors are selected by stepwise discriminant function. The order of selection in best subset logistic regression is different from other procedures. But in both stepwise logistic regression and stepwise discriminant analysis, the four predictors UNCONS, PCD, CRE and SPO2 are entered in the same order. Although the assumptions of discriminant analysis are violated the, direction of the relationship with ICU mortality are same. The estimated coefficient of best subset logistic model and stepwise logistic model are the same because same predictors are selected by both procedure and overall larger coefficients are estimated by these two procedures. These two methods were also compared by Pohar *et al.* (2004) using several simulated datasets. The authors concluded that linear discriminant analysis is a more appropriate method when the explanatory variables are normally distributed. For categorized predictor variables, linear discriminant analysis remains preferable, and logistic regression overcomes discriminant analysis only when the number of categories is small (2 or 3). When the assumptions of linear discriminant analysis are not met, the usage of it is not justified, while logistic regression gives good results regardless of the distribution of the predictors. Another recent study conducted by George *et al.* (2009) applied both logistic regression and discriminant analyses and concluded that logistic and discriminant analyses resulted in similar models, and although linear discriminant function is a better method than logistic regression when the normality assumptions are met. The performances of the selected models were evaluated by assessing the correct classification rate (CCR) and misclassification rate (MCR). The sensitivity of both the best subset logistic model and stepwise logistic model was 97.7% which showed that 98% died patients are correctly identified as death outcome and 2% are incorrectly identified as survived outcome. The specificity of this model is 31.4% which implies that 31% survived patients correctly identified as survived outcome. For both models, the over all CCR was 90%, which was better than chance (50%) and this CCR indicates that the predictors selected in these models, discriminate well between those who died and lived. The sensitivity of stepwise discriminant model was 81.9%, indicating that 82% died patients are correctly identified as death outcome and specificity of 71.4% indicates that 71.4 survived patients are correctly identified as survived outcome.

Logistic Discrimination

Fisher's linear discriminant function is inappropriate when most of the predictors are qualitative or when the assumption of multivariate normality is not met. In this case, the LDF can be estimated by maximum likelihood method. Now the LDF are estimated by maximum likelihood method by using BMDP-LE. The maximum likelihood estimates of the parameters and standard errors are given in table 3. The estimated co-efficient of first four predictors and for TEMP are close to the coefficient of logistic regression because this method also involves the maximum likelihood estimation. The maximum likelihood estimate of MV is not significant at 5% level of significance, the small difference in estimated coefficients of logistic discrimination and logistic model is due to the inclusion of insignificant predictor MV. After deleting MV, the estimates of logistic discrimination will become exactly equal to the estimates of logistic model. From the maximum likelihood estimate of LDF, the *classification rule* becomes, allocate patient with UNCONS, PCD, CRE, SPO2 and TEMP to group died if

$$-26.6751 + 2.3495UNCONS + 1.6975PCD + 2.382CRE - 1.5892SPO2 + 0.2771TEMP > 0$$

The overall CCR for stepwise discriminant model is 80.7% and overall CCR for logistic discrimination is 90%, which is about 9.3% higher. This implies that logistic discrimination improves the classification accuracy. The discriminating ability of logistic discrimination model with best subset and stepwise logistic model are also compared by ROC curve shown in the following **(Table 3- Fig. 3).**

Table 3- Maximum Likelihood Estimates of Linear Discriminant Function

Parameter	Estimate	Standard Error	Significance
Intercept	-24.5369	11.9211	0.0400
UNCONS	2.2644	0.4554	0.0000
PCD	1.7031	0.4621	0.0000
CRE	2.3973	0.8083	0.0030
SPO2	-1.6738	0.6425	0.0090
MV	0.6042	0.4465	0.1760
TEMP	0.2528	0.1201	0.0350

5. CONCLUSION

The three procedures are compared with respect to classification accuracy and discriminating ability and concluded that logistic regression models had greater power than discriminant model. Thus when the assumptions of discriminant analysis are violated, and the purpose of the study is classification, then its use is not misleading in terms of the variables selected, or the signs of the coefficients but magnitude of the coefficients, may be biased. So in this case, the coefficients of linear discriminant function must be estimated by maximum likelihood approach called Logistic discrimination. This study showed logistic discrimination improves the classification and discriminating ability of the discriminant analysis and provided the same results as obtained by logistic regression. Thus stepwise discriminant model can be used as a variable selection technique, but coefficients must be estimated by logistic models.

The risk factors of ICU mortality in this study are different from international studies because facilities in ICU's of Pakistan are challenging and critical care medicine is not yet recognized as a specialty is Pakistan. Comparing to western countries, patients are younger and often suffer from more reversible diseases. If they given a chance to go through, they can come back to socially and economically active life. A couple of high profile institutions need to take the lead and a national strategy has to be designated to improve ICU facilities for critically ill patients.

REFERENCE:

Ashish, N., Dilip, R. K., Sudarshan, S., Lucila, O., and Isaac, K. (2004) "Predicting mortality in an Indian Intensive Care." *Intensive Care Medicine*, Vol. **(30):** 2, 248-253.

Fabian, J., F. Jorge, A. Diego, and M. Carlos, (2005) "Comparison between logistic regression and neural networks to predict death in patients with suspected sepsis in the emergency room." *Journal of Critical Care*, Vol. **(9):** (2), R150-R156.

Chin, S. (2001)"The rise and fall of logistic regression." *Australian Epidemiology*, Vol. **(8):**7-10.

Gavilan, J., Gavilan, C. Manos-Pujal, M., and Herranz, J. (2007) "Discriminant analysis in predicting survival of patients with cancer of the larynx or hypopharynx." *Clinical Otolaryngology & Allied Sciences*, Vol. **(12):** (5), 331-335.

George, A., B. P. Demosthenes, N. P. Kostas, and T. Anastasia, (2009) "Logistic Regression and Linear Discriminant Analyses in Evaluating Factors Associated with Asthma Prevalence among 10- to 12-Years-Old Children: Divergence and Similarity of the Two Statistical Methods." *International Journalof Pediatrics*,

Hosmer, D. W., B. Javanovic, and S. Lemeshow, (1989). "Best Subsets Logistic Regression." *Biometrics*, Vol. **(45):** (4), 1265-1270.

Hosmer, D. W and S. Lemeshow, (2000) "*Applied Logistic Regression*." 2nd edition. New York, USA: John Wiley and Sons.

Kellie, J. A., S. Lemeshow, and D. W. Hosmer, (2007) "Goodness-of-fit tests for logistic regression models when data are collected using a complex sampling design." *Journal of computational Statistics & Data Analysis*, Vol. **(51):** (9), 4450-4464.

Lemeshow, S., D. Teres, and H. Pastides, (1988) "Predicting the Outcome of Intensive Care Unit Patients". *Journal of the American Statistical Association*, Vol. **(83):** (402), 348-356.

Levinson, S. S. (2007) "Comparison of Apolipoprotein B and Non-High- Density Lipoprotein Cholesterol." *American Journal of Clinical Pathology*, Vol. **(127):** (3), 449-455..

Mickey, J., and S. Greenland, (1989). "A study of impact of confounder selection vriteria on effect estimation." *American Journal of Epidemiology*, Vol. **(125):** 125-137.

Nele, V.D.N., V. Dirk, A. Marcel, and C. Francis, (1999) "Intensive Care for Very Eldery patients: Outcome and Risk Factors for in Hospital Mortality." *Journal of Age and Ageing*, Vol. **(28):**.253-256.

Pohar M., M. Blas, and S. Turk, (2004). "Comparison of logistic regression and linear discriminant analysis: a simulation study." *Metodoloˇski Zvezki,* Vol. **(1):** (1), 143-161.

Shahadat, S. H. and A. Farzana (2005) "Discriminant analysis and logistic discrimination." *Institute of Statistical Research and Training, university of Dhaka*. Abstract of MS thesis. Batch 3.