International Journal of Advanced Medical Sciences and Technology (IJAMST)
ISSN: 2582-7596 (Online), Volume-1 Issue-1, December 2020

Applying the Binary Logistic Regression Model for Prediction of the Ischemic Heart Disease

Chaithra N

Abstract: ***Background: Ischemic Heart Disease (IHD) has been taxing the healthcare systems with a huge economic burden by being the major cause of deaths globally. In that, India contributes to about one-fifth of such deaths. Statistics suggest that for every 1000 people in a population of at least 7.5 people suffer from this condition with an average age ranging between 30 and 69 years. In the present study, we evaluated individual contributions of M-Mode two dimensional echocardiographic parameters to determine the presence of IHD in a large segment of the Indian population using a logistic model. This model can be a predictive step in assisting the junior cardiologists/echo technicians to diagnose IHD patients well in advance and model could be used in software applications for the medical field and estimating the impact of health interventions in developing countries. Methodology: A total of 7304 echo records were selected for performing the logistic regression from Electronic Health Records (EHRs) at the Department of Cardiology, JSS Hospital. The data set included 6191 patients without IHD and 1113 patients with IHD. The study included one dichotomous variable and fifteen explanatory variables that were taken during the transthoracic echocardiography examination.log-likelihood Statistic, Cox and Snell R^2, Nagelkerke R^2, Akaike Information Criterion are the tests to find the Goodness of fit for testing the fitness of the model. We used Likelihood ratio and Wald tests for testing the statistical significance of regression co-efficient. The classification table and Receiver Operating Characteristic (ROC) curve are the method to evaluate the predictive accuracy of the logistic regression. Results: This study is the first to apply a large sample from echo data, to determine how well a predictive model would perform based only upon patients M-mode echocardiography measurements without clinical risk factors or physical exam findings. All the variables exhibited statistically significant variation between IHD patients and non-IHD patients. The prevalence of IHD was significantly higher in men than in women. Our model was constructed by a Likelihood ratio forward method and Iteration History shows that estimation was terminated at iteration number 8 with 9 Steps (Model) because the parameter estimates did not change by more than 0.001. Conclusion: The present study estimates the efficiency of the logistic model to investigate the factors contributing significantly to enhancing the risk of IHD and the resulting model has a higher accuracy rate (96.7%), which makes it a handy tool for junior cardiologists and echo technicians to screen the patients who have a high probability of having the disease and transfer those patients to senior cardiologists for further clinical evaluation.***

Keywords: Ischemic Heart Disease, Echocardiographic data, Goodness of fit, Prediction model indexes, Logistic regression

Manuscript received on 29 October 2020 | Revised Manuscript received on 07 November 2021 | Manuscript Accepted on 15 December 2020 | Manuscript published on 30 December 2020.

***** Correspondence Author

Ms. Chaithra N*, Assistant Professor, Division of Medical Statistics, Faculty of Life Sciences, JSS Academy of Higher Education &Research, Mysuru-15, Karnataka, India. E-mail: chaithra.mstats@jssuni.edu.in

I. INTRODUCTION

Ischemic heart disease (IHD) is a condition characterized by inadequate myocardial perfusion caused by reduced blood supply, increased myocardial oxygen demand, or both [1].This can be a partial or complete blockage mainly caused by a build-up of plaque (fatty deposits, in a process known as atherosclerosis) on the artery walls. The hardening of this plaque over time narrows the coronary artery and reduces blood supply to the heart. This is in turn augmented by the rupturing of platelets that clump together forming blood clots. These clots stick to the walls of the artery and further narrow down the vessels. The resultant of this blockage leads to a low supply of oxygen along with a reduction in the delivery of nutrients to the heart muscles thereby reducing the normal functioning of the heart[2,3].IHD has been taxing the healthcare systems with a huge economic burden by being the major cause of deaths globally. In that, India contributes to about one-fifth of such deaths [4]. Therefore, it is the need of the hour to manage this chronic disorder at the country level in order to curb the increasing global cardiovascular mortality [5]. Statistics suggest that among the overall cardiovascular deaths, 0.9 million (68.4%) is caused by IHDs and is increasing to a greater number in the years to come. Presently, for every 1000 people in a population of at least 7.5 people suffer from this condition with an average age ranging between 30 and 69 years. Although this appears to be a broad range, a recent observation is that the population in the low and middle-income groups with IHDs die at a young age [6]. Urbanization has been the greatest contributor to this disease due to a drastic lifestyle change among the population, whereas the rural population developing IHD is relatively lower but not nil [7]. The total IHD deaths and mortality rates with respect to different age groups need to be determined in order to optimize an effective management program. With the presently available therapies failing to optimally address the problems associated with IHDs, it has augmented to the challenges faced by the cardiovascular physicians in treating the condition. It has been observed that targeting the low and middle-income class of people to design management strategies has been an effective way to address this situation ona global scale [8].Logistic regression is a type of predictive model used for statistically predicting the outcome of a categorical dependent variable from a set of independent variables[9,10]. In medical research, it is used to generate models from which predictions can be made about the likelihood that an IHD is present or absent. In the present study, we evaluated individual contributions of M-Mode Two Dimensional (2D) echocardiographic (echo) parameters to determine the presence of IHD in a large segment of the Indian population using a logistic model [11].

In this regard, a database containing various parameters influencing the ischemic heart disease from patients can be prepared, which will further assist in predicting the same for the new patients using the logistic regression model. This can be a predictive step in assisting the junior cardiologists/echo technicians to diagnose IHD patients well in advance. The original motivation for this study was to determine if a clinical software application could be written using the binary logistic regression model that could successfully predict the likelihood of ischemic heart disease and this type of prediction model could be used in software applications for medical field and estimating the impact of health interventions in developingcountries[12].

II. DATA ANDMETHODOLOGY

2.1 Study subjects and data set

A retrospective analysis was performed atthe Department of Cardiology, JSS Hospital, Mysore. A total of 7304echo records were selected for performing the logistic regression from Electronic Health Records(EHRs).The data set included 6191 patients without IHD and 1113 patients with IHD. The study includedone dichotomous variable and fifteen explanatory variables that were taken during the transthoracic echocardiography examination. The dependent variable of the modelwas the presence or absence of the diseasein each patient and the independent variables were – Age, Aortic Root (AO), Left Atrium (LA), Right Ventricle (RV), Left Ventricle Internal Diameter during Diastole (LVID_d),LeftVentricleInternalDiastoleduringSystole (LVID_s),Intact Ventricular Septum Diameter during Diastole(IVS_d),Intact Ventricular Septum Diastole during Systole(IVS_s),Left Ventricular Posterior Wall Diameter during Diastole(LVPW_d), Left Ventricular Posterior Wall Diastole during Systole (LVPW_s),End Diastolic Volume (EDV), End Systolic Volume (ESV),Stroke Volume (SV), Ejection Fraction (EF (%)), and fractional Short (FS(%)).The study included one dichotomous variable and fifteen explanatory variables that were taken during the transthoracic echocardiography examination.log-likelihood Statistic, Cox and Snell R^2, Nagelkerke R^2, Akaike Information Criterion are the tests to find the Goodness of fit for testing the fitness of the model. We used Likelihood ratio and Wald tests for testing the statistical significance of regression co-efficient. The classification table and Receiver Operating Characteristic (ROC) curve are the method to evaluate the predictive accuracy of the logistic regression.The analysis was carried out using Statistical Package for the Social Sciences (IBM SPSS Statistics 22.0) and prediction model indexes were measured by STAT CRAFTsoftware.

2.1.1 Inclusion criteria

The following datasets of individuals were included:

- Aged between 18 and 96years.
- Normal adultpatients.
- Ischemic Heart Diseases,Ischemic Cardiomyopathy.
- LV DiastolicDysfunction.
- LV Systolicfunction
- Regional Wall Motion Abnormality.
- Myocardialinfarction

2.1.2 Exclusion criteria

The following cardiovascular disease were excluded from participation in the study as these could have influenced their echocardiography data:

- Aortic ValveSclerosis.
- Congenital HeartDisease.
- Concentric Left VentricleHypertrophy.
- Rheumatic HeartDisease.
- Degenerative Aortic, Mitral Valve Diseaseand other Cardiovasculardiseases

2.2. Goodness-Of-Fit

Goodness of fit is considered imperative in terms of logistic regression analysis for testing the fitness of the model. These tests are proposed mainly for overall measures of fit.The null hypothesis suggests that the model is correct in all aspects whereas the alternative hypothesis suggests the lack of fitness in themodel.

1) Pearson's Chi-square and Deviancetest

Pearson's chi-square test and the Deviance test are two goodness-of-fit tests obtained by measuring the difference between the observed dependent variable y and its fitted values and the data as described by a J ×2 contingency table, where the J rows are defined by the number of possible distinct values of covariate vector *X* and 2 columns are defined by the values of the Binary outcome variable y.

$$\chi^2 = \sum_{i=1}^{J} (y_i - m_i\pi_i)/m_i\pi_i(1-\pi_i)$$

$$D = 2\{\sum_{i=1}^{J} (y_i \ln(y_i/m_i\pi_i) + (m_i - y_i)\ln(m_i - y_i/(1-\pi_i))\}$$

where $y_i$is the observed dependent variable for the ith value and $\pi_i$is the fitted value for covariate pattern x_j.

Deviance is also called as -2 Log-likelihood (-2LL), which is indicative of the amount of unexplained information available after the fitness of the model. Therefore, large values of the log-likelihood statistic indicate poorly fitted statistical models, because larger the value of log-likelihood, more is the unexplainedobservations[13,14]

2) R^2 for LogisticRegression

In linear regression, the coefficient of determination is called R squared, it represents how much of the variance in the binary variable can be explained by the independent variable.

The cox and Snell's R^2_{cs}, which is based on the deviance of the model (-2LL (new)) and the deviance of the original model (-2LL (baseline)), and the sample size, n. It can be interpreted like R^2 in a multiple regression, but cannot reach a maximum value of 1. The Nagelkerke R2 can reach a maximum of 1.

Cox and Snell's R^2 is calculated from this equation:

$$R^2_{cs} = 1\text{-}\exp^{\frac{(-2LLnew-(-2LL(baseline)}{n}}$$

Nagelkerke's adjustment is calculated from:

$$R_N^2 = \frac{R_{cs}^2}{1 - exp - \frac{-2LLbaseline}{n}}$$

Although all of these measures differ in their computation, conceptually they are somewhat the same. So, in terms of interpretation, they can be similar to R^2 in linear regression because they provide a gauge forthe substantive significance of the model [13,15, 16]

3) Akaike Information Criterion (AIC)value

Akaike Information Criterion (AIC) is useful when we have more than one model to compare the goodness of fit. It is a maximum likelihood estimate which penalizes to prevent overfitting. It measures the flexibility of the models. A good model is the one that has minimum AIC among all the other models. The following equations are used to estimate the AIC of a model:

AIC = -2 ln (L) + 2 *k

where *L* is the value of likelihood and *k* is the number of estimated parameters [17,18].

2.3 Testing the statistical significance of regression co-efficient

As in linear regression, we want to know not only how well the model overall fits the data, but also the individual contribution of predictors.Here, we commonly used Likelihood ratio and Wald tests for testing the overall significance of the logistic regression model.

1) Likelihood ratio test

The LR test used to access the overall model fit can also be used for testing the significance of the individual regression coefficient. The distribution of the LR statistics is closely approximated by the Chi-square distribution for large samplesizes.

The comparison of observed to predicted values using the likelihood function is based on the following expression.

$$\text{LR} = -2 \log \frac{likelihoodofthefittedmodel}{likelihoodofthesaturatedmodel}$$

Assuming that there are *r* variables in the model under consideration which is the fitted model and saturated model is fitting a regression model when there are only two data points.Usingminustwiceitslogisnecessarytoobtaina quantity whose distribution is known and can, therefore,be used for hypothesis testing purposes.Such atest is called the likelihood ratio test. If LR statistics is less than 0.05,we can conclude that at least one of the exploratory variables contribute to the prediction of the outcome [19].

2) Wald statistic

The Wald-statistic is used to ascertain whether a variable is asignificantpredictoroftheoutcomeanditwasdeveloped by Abraham Wald. The test statistic is calculated as follows:

$$Z = \frac{i \ \beta^2}{SE\beta_j^2}$$

If the individual coefficient for the predictor is significantly different from zero, it is assumed that a significant contribution is made by the predictor in the case of z statistics. In such a case, the Wald test is applied, which is the ratio of the square of the regression coefficient to the square of the standard error of the coefficient confidence interval for the regression. The formula for the limits of a 100 (1-α) % two-sided confidence interval is[13, 20, 21].

$$\beta_j \pm |Z_{\alpha/2}| \, SE_{2_{\beta_j}}$$

where β_j is the coefficient of the estimate of the parameter and $SE_{\beta_j}^2$ is the standarderror of the estimate.

2.4. Predictive Accuracy of the Model

1) Classification table

The classification table is a method to evaluate the predictive accuracy of the logistic regression model. In this table, the observed values for the dependent outcome and the predicted values are cross-classified with two rows and two columns that report the number of the four outcomes of a binary classifier usually denoted as True Positive (a), False Negative (b), False Positive (c) and True Negative (d). If the logistic regression model has a good fit, we expect to see many counts in the *a* and *d* cells, and few in the *b* and *c* cells. Higher sensitivity and specificity indicate a better fit of the model.

Table 1: Classification Table

Observed	Predicted	
	1 (Yes)	0 (No)
1 (Yes)	True Positive (a)	False Negative (b)
0 (No)	False Positive (c)	True Negative (d)

Critical terms associated with classification table as follows:

Sensitivity

Sensitivity is the ability to assess accurately the people with the disease in a population (a + b), which determines the proportion of people with disease. Sensitivity as a fixed test characteristic provides a **True Positive Rate (TPR)**. A test with 100% sensitivity will recognize all patients with the disease by testing positive.

Mathematically, this can be expressed as:

$$\textbf{Sensitivity} = \frac{\textbf{TruePositive}}{\textbf{TruePositive+FalseNegative}} = \frac{a}{a+b}$$

Specificity

The specificity of a test also referred to as the **True Negative Rate (TNR)**, is the proportion of negative cases that were classified correctly. It describes the accuracy of the test to detection-diseased (d) individuals (c + d).The following equation is used to calculate a test'ssensitivity:

$$\textbf{Specificity} = \frac{\textbf{TrueNegative}}{\textbf{FalsePositive+TrueNegative}} = \frac{d}{c+d}$$

False Negative (b): Predicted no, but they actually do have the disease. (Also known as a "Type II error.")

False Positive(c): Predicted yes, but they don't actually have the disease. (Also known as a "Type I error.")

Predicted value of a positive test: Probability (disease present given that the test result ispositive).

PV+ = $\frac{a}{a+c}$

Predicted value of a negative test: Probability (disease absent given that the test result is negative).

PV- = $\frac{d}{b+d}$

The Accuracy (AC): Overall probability that a patient will be correctly classified. It is determined using the equation.

Accuracy = $\frac{\mathbf{TP+TN}}{\mathbf{TP+TN+FP+FN}} = \frac{(a+b)}{(a+d+c+b)}$

where TP – True Positive, TN – True Negative, FP – False Positive and FN – False Negative

Precision: It is defined as the proportion of the positive cases that had been predicted were correct.

Precision = $\frac{\mathbf{TP}}{\mathbf{TP+FP}}$

F-score: It is a measure of test accuracy and is the harmonic average of the precision p and recall r.

F-score = 2 x $\frac{precisionXrecall}{precision+recall}$

where recall is the number of correct positive results divided by the number of all relevant samples.

Kappa statistic: It measures the agreement between two raters who each classify N items into C mutually exclusive categories.

K = $\frac{p_o - p_e}{1-p_e}$

where p_o is the relative observed agreement among raters and p_e is the hypothetical probability of chance agreement [22-25].

2) Odds Ratios(OR)

Odds Ratios (OR) is a ratio of the odds of having a disease (or event) in exposed (high risk) versus non-exposed (low risk) groups.It compares the two odds relative to different events. For two events A and B, the odds of event A are defined as the probability that A does happen divided by the probability that it does not happen and can be estimated by logistic regression [26]:

$$\text{Odds} = \frac{\text{P (A happens)}}{\text{P (A does nothappen)}} = \frac{\text{P (A)}}{1 - \text{P(A)}}$$

When a logistic regression is calculated, the regression coefficient (β_1) is the estimated increase in the logged odds of the outcome per unit increase in the value of the independent variable. In other words, the exponential function of the regression coefficient ($e^{\beta 1}$) is the OR associated with a one-unit increase in the independent variable. The OR can also be used to determine whether a particular exposure is a risk factor for a particular outcome, and to compare the magnitude of various risk factors for that outcome. The OR value of 1 indicates no effect of the exposure on the outcome, whereas OR>1 indicates exposure associated with higher odds of the outcome and OR<1 indicates exposure associated with lower odds of the outcome [26, 27].

3) Receiver Operating Characteristic (ROC) curve.

The plot of sensitivity versus Specificity is called the receiver operating characteristic (ROC) curve and the area under the ROC curve (AUC) is a measure of how well a parameter can distinguish between two diagnostic groups (diseased/normal). This curve plays a central role in evaluating the diagnostic ability of tests to discriminate the truestateofsubjects.Forlogisticregressionanalyses,after fitting a model, each subject's fitted response probability π can be calculated. Using these probabilities as values of a separator, we can construct a nonparametric ROC curve by plotting the true positive rate (sensitivity) on the y-axis against the false-positive rate (1-specificity) on the x-axis. Accuracy is measured by the area under the ROC curve, which varies from 0.5 (no predictive ability) to1.0 (perfect predictive ability). An area of 1 signifiesperfect classification accuracy and AUC = 0.5 representpoor classification results.The c statistics or AUC represents the proportion of the subject pairs with different observed outcomes.The model correctly predicts a higher probability for observations with the event outcome than the probability for non-event observations.The c statistics range from 0.5 to 1 wherein0.5 value infersthat the model merely assigns observations randomly into the outcome group whereas the value of 1 determines a higher probability ofall observations with the event outcome, compared with that of the non-event observations. If several models were fitted to the same data set, the model chosen as the best model should be associated with the highest c statistic. Thus, the c statistic provides a basis for comparing different models fitted to the same data or the same model fitted to different data sets [28,29].

III. RESULTS ANDDISCUSSION

The purpose of this study was to develop a diagnosis model that performs well in predicting the presence of IHD. A total of 7304 cases with no missing values were selected for the analysis in that 1113 patients were diagnosed with IHD and 6191 normal patients were categorized as the study subject. All the variables exhibited statistically significant variation between IHD patients and non-IHD patients. The prevalence of IHD was significantly higher in men than in women. This study is the first to apply a large sample from echo data, to determine how well a predictive model would perform based only upon patients M-mode echocardiography measurements without clinical risk factors or physical exam findings. When the clinical evaluation is complete, the practitioner must determine whether the probability of IHD is sufficient to recommend further testing, which is often a standard exercise test. When the probability of disease is <5%, further testing is usually not warranted because the likelihood of a true negative rate is actually higher than that of a true positive rate. From this study, our results lead us to believe that such a prediction model can assist the physician in making accurate diagnosis well in advanceand useful in making decisions relating to the diagnosis of IHD [11,12, 29, 30].

3.1 Building amodel

The data which contains 1 dichotomous categorical outcome variable (y) and 15 predictor variables (x1, x2, x3, x4, …, x15), the outcome variable was coded 1 (event occurred) and 0 (event did not occur). Our model was constructed by a Likelihood ratio forward method, which was used to establish a mathematical model of the correlation between the variables andIHD.

The logistic regression procedure maximised its predictions of category membership by a highly computer-intensive process, which generates successive approximations called iterations and the model uses an iterative maximum likelihood algorithm to fit the data. The algorithm looks around to see if the fit would be improved by using different estimates. If it improves then it moves in that direction and then fits the model again. The algorithm stops when no significant additional improvement can be done. "Number of Fisher Scoring iterations," tells "how many iterations this algorithm run before it stopped". Here it is 8 Fisher Scoring iteration steps involved to fit themodel.

Table 2describes History for Block 1 (Forward: LR): method selection allows to specify the manner in which independent variables are entered into the analysis *Forward Selection (Likelihood Ratio)*. The model is improved stepwise by adding an extra variable to the model at each step. The variable which shows the"most significant" effect when correcting for the other variables in the model is added. The process stops if the step does not show a further improvement of the model after such corrections.The Iteration History shows that estimation was terminated at iteration number 8 with 9 Steps (Model) because the parameter estimates did not change by more than 0.001. At each iteration, the -2 log-likelihood decreases because it represents the unexplained variance in the outcome variable. Therefore, the smaller the value, better the fit in themodel.

Table 2. Iteration Historyfor Block 1: Method = Forward Stepwise (Likelihood Ratio).

Iteration		-2 Log likelihood	Coefficients									
			Constant	EF	IVS_S	Age	EDV	LVPW_s	FS	LVPW_d	LA	IVS_d
Step 9	1	3158.301	6.748	-.166	.026	.002	.002	-.020	.026	.004	.006	.012
	2	1998.131	11.514	-.257	.055	.008	.004	-.045	-.013	.013	.017	.022
	3	1597.823	16.217	-.333	.078	.016	.007	-.066	-.088	.022	.034	.030
	4	1477.802	20.666	-.420	.096	.025	.011	-.087	-.126	.029	.050	.036
	5	1458.708	23.490	-.478	.107	.030	.013	-.102	-.138	.033	.057	.038
	6	1457.992	24.205	-.492	.111	.032	.013	-.106	-.140	.034	.058	.039
	7	1457.990	24.238	-.493	.111	.032	.013	-.106	-.140	.034	.058	.039
	8	1457.990	24.238	-.493	.111	.032	.013	-.106	-.140	.034	.058	.039

Table 3.Omnibus Tests of Model Coefficient.

		Chi-square	df	Sig.
Step 9	Step	3.348	1	.000
	Block	4776.984	9	.000
	Model	4776.984	9	.000

The Omnibus Tests of Model Coefficients table 3 reports the chi-square goodness of fit test, associated with each step in a stepwise model which indicates that all nine predictors contribute significantly to fit the model. Here, the model was built in 9 steps, by adding a predictor at each step. The omnibus test is for all the steps showing significant impact of M-mode two-dimensional echocardiography parameters on the prediction of IHD. Finally, we can conclude that the model has an adequate fit for the data with test statistic is $\chi 2 = 4776.984$ on 9 df, $p < 0.0001$ at 5% level of significance.

Table 4. Model Summary.

Model	-2 Log likelihood	AIC	Cox & Snell R Square	Nagelkerke R Square
1	1657.172	1687.172	0.466	0.811
2	1625.341	1655.341	0.468	0.815
3	1580.932	1610.932	0.471	0.821
4	1540.784	1570.784	0.474	0.826
5	1501.125	1531.125	0.477	0.831
6	1480.977	1510.977	0.478	0.833
7	1473.455	1503.455	0.479	0.834
8	1461.339	1491.339	0.48	0.836
9	1457.99	1487.99	0.48	0.836

Model Summary table4 displays the -2 Log likelihood, AIC, Cox& Snell and Nagelkerke R. The Akaike's Information Criteria (AIC) is used to compare different fitted models against each other. The AIC will take each model and rank them from best to worst. Table 4 shows AIC values for nine steps, namely, AIC1 = 1687.172, AIC2 = 1655.341, AIC3 = 1610.932, AIC4 = 1570.784, AIC5 = 1531.125, AIC6= 1510.977, AIC7 = 1503.455, AIC8 = 1491.339, AIC9= 1487.990. The model 9 gives the minimum AIC among all theothermodels.So,wecanconcludethatitisthebest model to fit the data. The values of Cox & Snell and Nagelkerke R square are sometimes referred to as *pseudo* R^2values, which summarizes the proportion of dichotomous variables associated with the explanatory variables. This indicates that 48% of the variation in the dependent variable has been explained by the model 9. The Cox &Snell R^2value will normally be lesser than the Nagelkerke R^2measure.

Here, theNagelkerkeR^2 value from the model summary table is 0.84 indicating that a strongassociation exists between the predictors and the prediction explained by the logistic model.

Table 5. Classification table.

Observed			Predicted		
			Diagnosis		
			No	Yes	Percentage Correct
Step 9	Diagnosis	No	6143	48	99.2
		Yes	194	919	82.6
	Overall Percentage				96.7

Logistic regression estimates the probability of an event (in this case, having Ischemic heart disease) occurring. If the estimated probability of the event occurring is greater than or equal to 0.5, SPSS Statistics classifies the event as occurring (Ischemic heart disease is present). On the contrary, if the probability is less than 0.5, SPSS Statistics classifies the event as not occurring (no Ischemic heart disease). Classification or prediction of cases using independent variables using logistic regression is routinely followed and therefore, it is imperative to check the effectiveness of the model against the actual classification. There are many methods to assess this with their usefulness often depending on the nature of the study conducted. However, all these methods revolve around the observed and predicted classifications, which are presented in the "**Classification Table**", as shown in table5. Here, the model correctly classifies 6143 people as not having IHD and 48 patients were predicted incorrectly as having IHD. This type of hypothesis testing is called Type I Error because 99.2% of true negative cases are correctly identified by the model. Similarly, 919 patients were correctly predicted to have diseasei.e., 82.5% of true positive cases were correctly predicted by the model and 194 patients were wrongly predicted as without IHD while they had the disease. So, it is very dangerous to predict a positive IHD patient as free from the disease. The overall accuracy of classification was correct 7062 out of 7304 times, for an overall success rate is 96.7% of the patients data were predicted perfectly by the logisticmodel.

An increasingly important issue in health care systems is cost-effectiveness. In general, the efficacy of any test system is based on its sensitivity, specificity, and prognostic value. However, the sensitivity of 0.825 (95 % CI = 0.80 to 0.84), indicates the probability of the test to correctly classify an individual as ′diseased′. When a test has high sensitivity, the maximum number of patients with the disease is picked up. Specificity of 0.993 (95 % CI = 0.98 to 0.99), refers to the probability that a test result will be negative when the disease is not present.Positive predictive value of0.952 (95 % CI = 0.93. to 0.96), indicates the probability that the disease is present when the test is positive and 0.969 (95 % CI = 0.96 to 0.97) suggests that the probability of a person without the disease when the test is negative. The model accuracy was good at 0.967 (95 % CI= 0.96 to 0.97) and precision was 0.969, which means the predicted IHD cases were correctly identified. These indexes infer prediction was good overall. Other indexes showed good outcomes: the model Kappa was 0.865, which indicated good prediction, while the F-score was0.981. In this study, the model prediction indexes were more than 0.8, which suggests that our refined model had good prediction performance in the prediction of IHD as shown in table6.

Table 6. Model prediction indexes for study subject.

Statistics by class	
Sensitivity	0.825 (95 % CI = 0.80 to 0.84)
Specificity	0.993 (95 % CI = 0.98 to 0.99)
Negative Predicted Value	0.969 (95 % CI = 0.96 to 0.97)
Positive Predicted Value	0.952 (95 % CI = 0.93. to 0.96)
Accuracy	0.967 (95 % CI = 0.96 to 0.97)
Precision	0.969
F-score	0.981
Kappa	0.865

Table 7. Variables in the equation.

		B	S.E.	Wald	df	Sig.	Exp(B)	95% C.I.for EXP(B)	
								Lower	Upper
Step 1	EF	-.559	.020	794.167	1	.000	.572	.550	.595
	Constant	29.593	1.128	687.747	1	.000			
Step 2	IVS_S	.097	.016	37.354	1	.000	1.101	1.068	1.136
	EF	-.564	.020	779.680	1	.000	.569	.547	.592
	Constant	28.588	1.148	620.397	1	.000			
Step 3	Age	.034	.005	42.413	1	.000	1.035	1.024	1.045
	IVS_S	.085	.016	26.634	1	.000	1.089	1.054	1.125
	EF	-.554	.020	739.715	1	.000	.575	.552	.598
	Constant	26.264	1.183	492.620	1	.000			
Step 4	Age	.036	.005	45.365	1	.000	1.037	1.026	1.048
	IVS_S	.063	.017	13.977	1	.000	1.065	1.031	1.101
	EDV	.016	.003	39.800	1	.000	1.016	1.011	1.021
	EF	-.543	.021	694.447	1	.000	.581	.558	.605
	Constant	24.351	1.214	402.566	1	.000			
Step 5	Age	.036	.005	45.654	1	.000	1.037	1.026	1.048
	IVS_S	.127	.026	23.973	1	.000	1.136	1.079	1.195
	LVPW_s	-.118	.015	58.056	1	.000	.889	.862	.916
	EDV	.017	.003	43.682	1	.000	1.017	1.012	1.022
	EF	-.561	.022	676.966	1	.000	.571	.547	.595
	Constant	25.973	1.298	400.488	1	.000			
Step 6	Age	.035	.005	42.768	1	.000	1.036	1.025	1.047
	IVS_S	.121	.025	23.140	1	.000	1.129	1.075	1.186
	LVPW_s	-.099	.015	43.721	1	.000	.905	.879	.933
	EDV	.017	.003	42.606	1	.000	1.017	1.012	1.022
	EF	-.496	.024	424.060	1	.000	.609	.581	.639
	FS	-.141	.028	24.709	1	.000	.868	.821	.918
	Constant	26.159	1.270	424.369	1	.000			
Step 7	Age	.035	.005	41.746	1	.000	1.036	1.025	1.047
	IVS_S	.118	.025	21.635	1	.000	1.126	1.071	1.183
	LVPW_d	.046	.013	12.943	1	.000	1.047	1.021	1.074
	LVPW_s	-.102	.015	47.831	1	.000	.903	.878	.930
	EDV	.017	.003	42.291	1	.000	1.017	1.012	1.022
	EF	-.495	.024	427.539	1	.000	.610	.582	.639
	FS	-.142	.028	25.380	1	.000	.867	.821	.917
	Constant	25.818	1.270	413.344	1	.000			
Step 8	Age	.032	.006	33.875	1	.000	1.033	1.022	1.044
	LA	.060	.017	12.026	1	.001	1.062	1.026	1.099
	IVS_S	.117	.026	20.898	1	.000	1.124	1.069	1.182
	LVPW_d	.044	.013	11.279	1	.001	1.045	1.019	1.073
	LVPW_s	-.105	.014	53.601	1	.000	.900	.875	.926
	EDV	.013	.003	23.386	1	.000	1.013	1.008	1.019
	EF	-.494	.024	428.980	1	.000	.610	.582	.639
	FS	-.138	.028	24.084	1	.000	.871	.824	.920
	Constant	24.343	1.328	336.235	1	.000			
Step 9	Age	.032	.006	33.219	1	.000	1.032	1.021	1.044
	LA	.058	.017	11.262	1	.001	1.060	1.025	1.097
	IVS_d	.039	.019	4.255	1	.039	1.040	1.002	1.079
	IVS_S	.111	.026	18.478	1	.000	1.117	1.062	1.175
	LVPW_d	.034	.016	4.483	1	.034	1.034	1.003	1.067

	LVPW_s	-.106	.014	54.723	1	.000	.900	.875	.925
	EDV	.013	.003	23.455	1	.000	1.013	1.008	1.019
	EF	-.493	.024	426.303	1	.000	.611	.583	.640
	FS	-.140	.028	24.370	1	.000	.869	.822	.919
	Constant	24.238	1.325	334.511	1	.000			

The final logistic model being tested is:

$$P(Y) = \frac{1}{1+e^{-(\beta 0 + \beta 1 X 1 + \beta 2 X 2 + \cdots + \beta n X n)}}$$

Model 1: EF Variable is entered on Step1:$P(Y) = \frac{1}{1+e^{-\beta 0 + \beta 1 X 1}} = \frac{1}{1+e^{-(29.593 - 0.559 * EF)}}$

Where β_0= 29.593, $\beta_1 X_1$ = -0.559(EF)

Model 2: IVS_S Variable is entered on Step 2:$P(Y) = \frac{1}{1+e^{-(\beta 0 + \beta 1 X 1 + \beta_2 X 2)}} = \frac{1}{1+e^{-(28.588 - 0.564 * EF + 0.097 * IVS\_S)}}$

Where β_0 = 28.588, $\beta_1 X_1$= -0.564(EF), $\beta_2 X_2$ = 0.097(IVS_S)

Model 3: Age Variable is entered on Step 3:

$P(Y) = \frac{1}{1+e^{-\beta 0 + \beta 1 X 1 + \beta_2 X_2 + \beta_3 X_3}} = \frac{1}{1+e^{-(26.264 - 0.554 * EF + 0.085 * IVS\_S + 0.034 * Age)}}$

Where β_0 = 26.264, $\beta_1 X_1$= - 0.554(EF), $\beta_2 X_2$ = 0.085(IVS_S), $\beta_3 X_3$ = 0.034(Age)

Model 4: EDV Variable is entered on Step 4:

$P(Y) = \frac{1}{1+e^{-\beta 0 + \beta 1 X 1 + \beta_2 X 2 + \beta_3 X 3 + \beta_4 X 4}} = \frac{1}{1+e^{-(24.351 - 0.543 * EF + 0.063 * IVS\_S + 0.036 * Age + 0.016 * EDV)}}$

Where β_0 = 24.351, $\beta_1 X_1$= -0.543(EF), $\beta_2 X_2$ = 0.063(IVS_S), $\beta_3 X_3$ = 0.036(Age), $\beta_4 X_4$ = 0.016(EDV)

Model 5: LVPW_s Variable is entered on Step 5:

$P(Y) = \frac{1}{1+e^{-\beta 0 + \beta 1 X 1 + \beta_2 X 2 + \beta_3 X 3 + \beta_4 X 4 + \beta_5 X 5}} = \frac{1}{1+e^{-(25.073 - 0.561 * EF + 0.127 * IVS\_S + 0.036 * Age + 0.017 * EDV - 0.118 * LVPW\_s)}}$

Where β_0 = 25.073, $\beta_1 X_1$= -0.561(EF), $\beta_2 X_2$ = 0.127(IVS_S), $\beta_3 X_3$ = 0.036(Age, $\beta_4 X_4$ = 0.017(EDV), $\beta_5 X_5$ = -0.118(LVPW_s)

Model 6: FS Variable is entered on Step 6:

$P(Y) = \frac{1}{3+\beta_4 X 4 + \beta_5 X 5 + \beta_6 X 6} 1 + e^{-\beta 0 + \beta 1 X 1 + \beta_2 X 2 + \beta_3 X}$

$= \frac{1}{1+e^{-(26.159 - 0.496 * EF + 0.121 * IVS\_S + 0.035 * Age + 0.017 * EDV - 0.099 * LVPW\_s - 0.496 * FS)}}$

Where β_0 = 26.159, $\beta_1 X_1$= -0.496(EF), $\beta_2 X_2$ = 0.121(IVS_S), $\beta_3 X_3$ = 0.035(Age), $\beta_4 X_4$ = 0.017(EDV), $\beta_5 X_5$ = -0.099(LVPW_s), $\beta_6 X_6$ = -0.496(FS)

Model 7: LVPW_d Variable is entered on Step 7:

$P(Y) = \frac{1}{4 X 4 + \beta_5 X 5 + \beta_6 X 6 + \beta_7 X 7} 1 + e^{-\beta 0 + \beta 1 X 1 + \beta_2 X 2 + \beta_3 X 3 + \beta}$

$= \frac{1}{1 + e^{-(25.818 - 0.495 * EF + 0.118 * IVS\_S + 0.035 * Age + 0.017 * EDV - 0.102 * LVPW\_s - 0.142 * FS + 0.046 * LVPW\_d)}}$

Where β_0 = 25.818, $\beta_1 X_1$= -0.495(EF), $\beta_2 X_2$ = 0.118(IVS_S), $\beta_3 X_3$ = 0.035(Age), $\beta_4 X_4$ = 0.017(EDV) $\beta_5 X_5$ = -0.012(LVPW_s), $\beta_6 X_6$ = -0.142(FS), $\beta_7 X_7$ = 0.046(LVPW_d)

Model 8:LA Variable is entered on Step 8:

$P(Y) = \frac{1}{4 + \beta_5 X 5 + \beta_6 X 6 + \beta_7 X 7 + \beta_8 X 8} 1 + e^{-\beta 0 + \beta 1 X 1 + \beta_2 X 2 + \beta_3 X 3 + \beta_4 X}$

$= \frac{1}{1+e^{-(24.343 - 0.494 * EF + 0.117 * IVS\_S + 0.032 * Age + 0.013 * EDV - 0.105 * LVPW\_s - 0.138 * FS + 0.044 * LVPW\_d + 0.060 * LA)}}$

Where β_0 = 24.343, β_1 X_1= -0.494(EF), β_2 X_2 = 0.117(IVS_S), β_3 X_3 = 0.032(Age), β_4 X_4 = 0.013(EDV)
β_5X_5 = -0.105(LVPW_s), β_6 X_6 = -0.138(FS), β_7 X_7 = 0.044(LVPW_d), β_8 X_8 = 0.060(LA)

Model 9: Final Model for prediction of IHD, IVS_d Variable is entered on Step 9

$$P(Y)=\frac{1}{1+e^{-\beta0+\beta1X1+\beta_2X2+\beta_3X3+\beta_4X4+\beta_5X5+\beta_6X6+\beta_7X7+\beta_8X8+\beta_9X9}}$$

$$=\frac{1}{1+e^{-(24.238—0.493*EF+0.111*IVS\_S+0.032*Age+0.013*EDV-0.106*LVPW\_s-0.140*FS+0.034*LVPW\_d+0.058*LA+0.039*IVS\_d)}}$$

Where β_0 = 24.238, β_1 X_1= -0.493(EF), β_2 X_2 = 0.111(IVS_S), β_3 X_3 = 0.032(Age), β_4 X_4 = 0.013(EDV), β_5 X_5 = -0.106(LVPW_s), β_6 X_6 = -0.140(FS), β_7 X_7 = 0.034(LVPW_d), β_8 X_8 = 0.058(LA), β_9 X_9 = 0.039(IVS_d)

Once the selection of the model (Step) has been completed, a final logistic model was constructed and the importance of each parameter included in the model was verified by an examination of the Wald statistic. In our study, the logistic regression model is fitted to the data using the selected variables and results are presented in Table 7. Variables in the equation table display the estimated regression coefficient, standard error (S.E), Wald statistic, df, Sig. (*p*-value); as well as the Exp(B) and confidence interval for the Exp(B). The results of the logistic regression analysis enabled us to determine which characteristics were independently associated with the presence ofIHD.

The Wald Statistic, which tests the hypothesis of whether the beta co-efficient for that predictor is significantly different from zero i.e, β = 0. If the coefficient significantly different from zero, we can assume that the independent variable is making a significant contribution to the prediction of the outcome (Y) and the size of the contribution of the nine factors is described by co-efficient values. For these data, it is observed that the independent variables: EF, LVPW_s, Age, FS, EDV, IVS_S, LA, LVPW_d and IVS_d are highly important factor in order to predict the risk of IHD and its p-value is less than 0.05 at 5% level of significance. The lower negativeregression coefficient value of EF, FS, and LVPW_s, higher the riskof the IHDdisease. Conversely, the positive regression co-efficient of age, EDV, IVS_S, LA, LVPW_d and IVS_dindicates the probability of disease risk increases with the increasing value of these factors. Note that Exp(B) is also known as the odds ratio predicted by themodel.

When used in logistic regression,the Odds Ratio (OR) determines the increase in the number of odds for an outcome per unit change in the associated *explanatory variable*. The odds ratio is labelled as **Exp(B)** on *SPSS*. Exp(B)istheexponentialvalueof*b*forthepredictorwith 95 % confidence interval for age: (OR =1.03, 95% CI =1.021-1.044),LA(OR=1.060,95%CI=1.025-1.097), IVS_d (OR = 1.040, 95% CI = 1.002 - 1.079), IVS_S (OR= 1.117,95%CI=1.062-1.175),LVPW_d(OR=1.034, 95%CI=1.003-1.067),LVPW_s(OR=0.900,95%CI= 0.875 – 0.925), EDV (OR = 1.013, 95% CI = 1.008-1.019),EF(OR=0.611(95%CI=0.583–0.640)andFS (OR = 0.869, 95% CI = 0.822 – 0.919).These indicate the interpretation for Exp(B). If the value is greater than one, then as the predictor increases, the odds of the outcome occurring increases. conversely, a value less than 1 indicates that as the predictor increases, the odds of the outcome occurring decreases. For the Interpretation to be reliable the confidence interval of exp(B) should not cross 1.The output from the current study suggested that the change in the coefficient of estimates from the sequential analysis was substantial. The critical evaluation of the individual predictor reveals that the selected variables in the final model are significantly associated with the response variable.

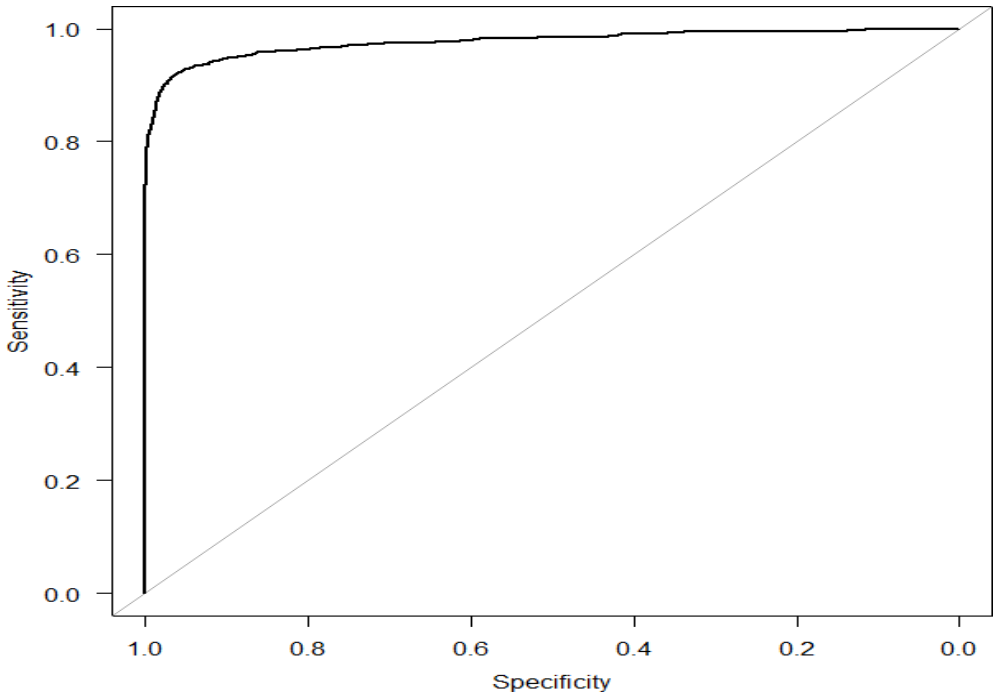

Figure1. The area under the Receiver Operating Characteristic (ROC) curve for model.

ROC analysis is used to quantify the accuracy of medical diagnostic tests to discriminate between two patient states, typically referred to as "diseased" and "no diseased". The AUC or the c-statistic is used to compare the goodness of fit of the logistic regression models. AUC values rangefrom 0.5 to 1.0. An AUC of 1.0 indicates a perfect test, whereas in general, an AUC of 0.9– 0.99 is an excellent test,0.8–0.89 a good test, 0.7–0.79 a fair test, 0.51–0.69 a poor test, and 0.5 is of no value. The data from all tested samples were used to generate these curves. The area under the ROC curve (AUC) for this logistic regression model for predicting IHD from M-Mode 2D echocardiographicparameterswas 0.976 with a 95 % class interval (0.97-0.98) indicating an excellent test to fit the model.

The main intention is to ensure cost-effective diagnosis as well as to improve clinical efficiency in healthcare systems. In addition, this can be developed as an automated computer program to predict the probability of IHD with our model, by inputting the results of M-mode echocardiographic measurement values into a computer. Further research is warranted for the development of more accurate strategies with the incorporation of personal and clinical characteristics with echocardiographic parameters to predict the risk of IHD using new statistical machine learning and data miningtechniques.

DISCLOSURE STATEMENT

No potential conflict of interest was reported by the authors.

IV. CONCLUSION

The present study estimates the efficiency of the logistic model to investigate the factors contributing significantly to enhancing the risk of IHDas well as accurately predict the overall risk. The above analysis shows that the model has a high specificity and sensitivity. Thus, the statistically significant factors are EF, LVPW_s, Age, FS, EDV, IVS_S, LA, LVPW_d, and IVS_d as expected by the logistic model, whereas the other factors are considered clinically important for the model. We considered 9 factors out of the15 from our variable group after the iteration process to predict the risk of IHD.The outcome of this study can be used for assisting cardiologists for accurate diagnosis of heart disease. Furthermore, the resulting model has a higher accuracy rate (96.7%), which makes it a handy tool for junior cardiologists and echo technicians to screen the patients who have a high probability of having the disease and transfer those patients to senior cardiologists for further clinicalevaluation.

REFERENCES

1. Falk E, Shah P, Feyter P. *Ischemic heart disease*. London: Manson Publishing;2007.
2. Steenbergen C, Frangogiannis NG. Ischemic Heart Disease. In: A.Hill J, Olsen EN. (eds.) *Muscle: Fundamental Biology and Mechanisms of Disease*. Elsevier Inc.; 2012. p.495–522.
3. Roncella A, Pristipino C. *Psychotherapy for ischemic heart disease: an evidence-based clinical approach*. Rome: Springer;2016.
4. Finegold JA, Asaria P, Francis DP. Mortality from ischaemic heart disease by country, region, and age : Statistics from World Health Organisation and United Nations. *International Journal* of Cardiology. 2013;168(2):934–45. Available from:http://dx.doi.org/10.1016/j.ijcard.2012.10.046.
5. Ke C, Gupta R, Xavier D, Prabhakaran D, Mathur P, Kalkonde Y V, et al. Divergent trends in ischaemic heart disease and stroke mortality in India from 2000 to 2015 : a nationally representative mortality study. *The Lancet Global Health*. 2018;6(8):914–23. Available from:DOI:10.1016/S2214-109X(18)30242-0.
6. Moran AE, Forouzanfar MH, Roth GA, Mensah GA, Ezzati M, Murray CJL. Temporal trends in ischemic heart disease mortality in 21 world regions, 1980 to2010 the global burden of disease 2010 study. *Circulation.* 2014;129(14):1483–92. from:DOI:10.1161/CIRCULATIONAHA.113.004042.Temporal.
7. Gupta R, Mohan I, Narula J. Trends in coronary heart disease epidemiology in India.*Annals of Global Health*. 2016;82(2):307–15. Available from:http://dx.doi.org/10.1016/j.aogh.2016.04.002.
8. Delgado R ED. Interventional treatment of advanced ischemic heart disease. London: Springer; 2009.
9. Bhatti IP,Lohano HD, Pirzado ZA, Jafri IA. A Logistic regression analysis of the ischemic heart disease risk.*Journal of Applied Sciences*. 2006;6(4):785–8. Available from:DOI:10.3923/jas.2006.785.788.
10. Iqbal F, Siddiqi AR, Sabir MA. Determining risk factors for ischemic heart disease using logistic regression and classification tree. *Sylwan*. 2014;158(6):69–87. Availablefrom:DOI:10.13140/2.1.1536.8329.
11. Xu H, Duan Z, Miao C, Geng S, Jin Y. Development of a diagnosis model for coronary artery disease.*Indian Heart Journal*. 2017;69(5):634–9. Available from:http://dx.doi.org/10.1016/j.ihj.2017.02.022.
12. Soleimani P, Neshati A. Applying the regression technique for prediction of the acute heart attack.*International Journal of Biomedical and Biological Engineering*.2015;9(11):767–71.
13. Field Andy. Discovering Statistics Using IBM SPSS Statistics.4th Edition. University of Sussex, UK: SAGE.2017
14. Bergerud, W.A. Introduction to Regression Models: with worked forestry examples. Biom. Info. Hand. 7. Res. Br., B.C. Min. For., Victoria, B.C. Work. Pap.26/1996.
15. Hilbe JM. *Practical Guide to Logistic Regression*. Boca Raton:Chapman and Hall/CRC; 2015. Available from: http://www.jstatsoft.org/v71/b03/.
16. Hosmer D, Lemeshow S. *Applied Logistic Regression*.Wiley Series in Probability and Statistics. 2nd ed. New York:Wiley-Interscience Publication;2000.
17. Wang Y, Liu Q. Comparison of Akaike Information Criterion (AIC) and Bayesian Information Criterion (BIC) in selection of stock- recruitment relationships.*Fisheries Research*. 2006;77:220–5. Availablefrom:DOI:10.1016/j.fishres.2005.08.011.
18. Menezes FS De, Liska GR, Cirillo MA, Vivanco MJF. Data classification with binary response through the Boosting algorithm and logistic regression. *Expert Systems with Applications*. 2017;69:62–73. Available from:http://dx.doi.org/10.1016/j.eswa.2016.08.014.
19. Ae H and Park. An introduction to logistic regression : from basic concepts to interpretation with particular attention to nursing domain. *Journal of Korean Academy of Nursing*.2013;43(2):154–64.
20. Kirkwood BR, Sterne JAC. *Essential Medical Statistics*. 2nd ed. Malden:Wiley-Blackwell; 2003. Available from: gums.ac.ir/Upload/modules/Contents/asset68/08654287191.pdf.
21. Elsalam NMMA. Binary logistic regression to identify the risk factors of eye glaucoma.*International Journal of Sciences: Basic and Applied Research (IJSBAR)*.2015;23(1):366–76.
22. Trevethan R. Sensitivity, specificity, and predictive values: foundations, pliabilities, and pitfalls in research and practice. *Frontiers in Public Health*. 2017 Nov 20;5:1–7. Available from: http://journal.frontiersin.org/article/10.3389/fpubh.2017.00307/full.
23. Gogtay NJ, Thatte UM. Statistical evaluation of diagnostic tests (part 1): sensitivity, specificity, positive and negative predictive values.*Journal of The Association of Physicians of India*. 2017;65:80– 4.
24. Giancristofaro RA, Salmaso L. Model performance analysis and modelvalidationinlogisticregression.*Statistica*.2003;63(2):375–96.
25. Baratloo A, Hosseini M, Negida A, El Ashal G. Part 1: Simple Definition and Calculation of Accuracy, Sensitivity, and Specificity. *Emergency*.2015;3(2):48–9.
26. Diaz-Quijano FA. A simple method for estimating relative risk using logistic regression. *BMC Medical Research Methodology*. 2012;12:14:1–6.
27. Sundaram KR, Dwivedi SN, Sreenivas V. *Medical Statistics:principles and methods*. 2nd ed. New Delhi: Wolters Kluwer; 2015.
28. Carter J V, Pan J, Rai SN, Galandiuk S. ROC-ing along: Evaluation and interpretation of receiver operating characteristic curves. *Surgery*. 2016;159(6):1638–45. Available from:http://dx.doi.org/10.1016/j.surg.2015.12.029.
29. Reychav I, Zhu L, McHaney R, Arbel Y. Empirical thresholding logistic regression model based on unbalanced cardiac patient data. *Procedia Computer Science*. 2017;121:160-165. Available from: https://doi.org/10.1016/j.procs.2017.11.022
30. Fihn SD, Gardin JM, Abrams J, Berra K, Blankenship J, Dallas A, et al. 2012 ACCF/AHA/ACP/AATS/PCNA/SCAI/STS guideline for the diagnosis and management of patients with stable ischemic heart disease: a report of the American College of Cardiology Foundation/American Heart Association Task Force on Practice Guidelines, and the American College of Physicians, American Association for Thoracic Surgery, Preventive Cardiovascular Nurses Association, Society for Cardiovascular Angiography and Interventions, and Society of Thoracic Surgeons. *Circulation*. 2012;126:354-471. Available from:https://doi.org/10.1161/CIR.0b013e318277d6a0.