Rangsit Journal of Arts and Sciences, January-June 2012 RJAS Vol. 2 No. 1, pp. 23-31

Multinomial logistic regression analysis of breast cancer

Doungporn Maiprasert[1]* and Krieng Kitbumrungrat[2]

[1,2]Faculty of Information Technology, Rangsit Univerisity, Patumthani 12000, Thailand
[1]E-mail: dmiprasert@yahoo.com; [2]E-mail: kr_stat@yahoo.com[2]

*Corresponding author

Submitted 5 February 2012; accepted in final form 6 June 2012

Abstract

This study aims at developing Multinomial Logistic Regression (MLR) to evaluate the probability of breast cancer, proposing MLR to predict five stages of breast cancer (Benign, I, II, III and IV). Nine characteristics of breast cancer: Clump Thickness (X_1); Uniformity of Cell Size (X_2); Uniformity of Cell Shape (X_3); Marginal Adhesion (X_4); Single Epithelial Cell Size (X_5); Bare Nuclei (X_6); Bland Chromatin (X_7); Normal Nucleoli (X_8); and Mitoses (X_9) are used as independent variables. Results show that Multinomial Logistic Regression (MLR) yields a coefficient of a model indicating that X_1 and X_6 have significance less than 0.05. Thus, the prediction of log – likelihood function for a classification staging of breast cancer with $P(Y \leq 4)$ of stage IV is a reference category, reducing a model as:

$$\log\left(\frac{p_i}{1-p_i}\right) = 819.992 + 608.852x_1 + 615.165x_6$$

Keywords**:** *multinomial logistic regression(MLR), logistic regression, classification, prediction*

1. Introduction

Breast cancer is the second leading cause of death for women, accounting for over 13% of all deaths. There are many different types of breast cancer with different stages including benign stage, stage I; II; III; and IV, respectively. Survival rates of breast cancer patients may increase when the disease is detected in the earlier stages of the disease through mammograms. The implementation of mass screening results in increased caseloads for radiologists, which may increase the chances of improper diagnosis. As such, the prediction using logistic regression will aid radiologists in detecting breast cancer.

MLR serving as a non-linear model on the dataset to select significant parameters through the 'Self-Consistency Test' is implemented to predict stages of breast cancer. This test is an examination for self-consistency of a prediction method. When the self-consistency test is performed, different features of breast cancer in the relevant dataset are then identified using the rule parameters derived from the same dataset which is the so-called "training dataset". Three different types of logistic regression analyses can be determined based on the type of scale in which the dependent variable is measured along with the category number of the dependent variable. In cases where the dependent variable is a categorical variable with two choices; it is called "Binary Logistic Regression Analysis" (Stephenson, 2008). For the dependent variable with only two categories or, in other words, a dichotomy, generally what is observed is the likelihood of an action to either occur or not is defined as the existence/absence of a feature (Long, 1997), which can be seen in a sample situation where a student's admission or non-admission in an academic program or a student with or without learning difficulty are indifferent. For the dependent variable, with more than two categories of classification, a "Multinomial Logistic Regression Analysis (MLR)" (Stephenson, 2008) can be applied. An example of which can be seen through a sample situation where a multinomial logistic regression analysis is applied for the estimation of a dependent variable consisting of students attending five different faculties. Having the dependent variable with more than two categories, polytomous, is a situation frequently. faced in application. However, the most important point to be considered here is whether the categories are ordinal or not since some models are only appropriate for ordinal categories, whereas the others can be used for categories with both ordinal and non-ordinal. If the dependent variable is obtained by ordinal scale and if the categories are ordinal, then "Ordinal Logistic Regression Analysis" is to be used

(Stephenson, 2008). This can be seen in a sample situation where a dependent variable with assertiveness levels of tests are grouped as "easy", "mid" and "high" requires the application of ordinal logistic regression analysis. As such, the dependent variable in a logistic regression analysis cannot be viewed as a continuous variable. However, the level of explanatory variables is not considered important, thus making it possible to yield mixed models.

2. Objectives

The objective of the present study of MLR is to predict the five stages of breast cancer (Benign, I, II, III and IV) considered as dependent variables, and to use nine characteristics of breast cancer: Clump Thickness (X_1); Uniformity of Cell Size(X_2); Uniformity of Cell Shape (X_3); Marginal Adhesion(X_4); Single Epithelial Cell Size(X_5); Bare Nuclei(X_6); Bland Chromatin(X_7); Normal Nucleoli (X_8); and Mitoses(X_9) as independent variables.

3. Methodology

3.1 Analysis of Multinomial Logistic Regression

MLR is used as a classification to predict the outcome of biopsy in breast cancer. The MLR is a generalization of the logistic regression model commonly used with the data comprising dependent variables known as "polytomous" and independent variables with numerical or categorical predictors.

The statistical test in MLR includes:

3.1.1 Chi – square is implemented to test these hypotheses:

H_0: The sample has been drawn from population following a specified distribution.

H_1: The sample has not been drawn from population following a specified distribution.

Chi-square test appropriates measures of agreement (or disagreement) between observed and expected frequencies. Chi-square is computed by dividing the squared difference between observed and expected frequencies in each set of frequencies by the expected frequency with the summation of the overall set. The interaction tests were performed to determine the significant values of each variable. The significance of the interaction is then measured and reported, the test is cross tabulated, and the values were determined by operating Pearson Chi-Square.

The Pearson Chi-Square is shown as follows:

$$X^2 = \sum_{i=1}^{r} \frac{(O_i - E_i)^2}{E_i} \quad (1)$$

where O_i is observed values

E_i is expected values

X^2 is chi-square value

If the X^2 value is more than the critical value, we reject the null hypothesis.

If the X^2 value is less than the critical value, we accept the null hypothesis.

3.1.2 Maximum likelihood estimate

The principle of maximum likelihood states that the use of estimation of β the value which maximizes the expression in this equation:

$$G^2 = -2[\ln L_p - \ln L_0]; df = p \quad (2)$$

where *Lp* is likelihood of constant value and group of independence *P*- value.

L_0 is likelihood of only constant value.

3.1.3 Relationship between independence value and dependence value (Wald test)

$$H_0 : \beta_i = 0$$
$$H_1 : \beta_i \neq 0$$

The Wald test statistic is the function of the difference of maximum likelihood estimate (MLE), and the value is hypothesized and normalized by an estimate of the standard deviation of the MLE. The equation (3) is shown as follows:

$$\chi^2 = \left[\frac{\hat{\beta}_i - \beta_i}{SE(\hat{\beta}_i)} \right]; df = p \quad (3)$$

where $SE(\hat{\beta}_i)$ is the standard of the maximum likelihood function, estimate is standard error and *df* is degree of freedom.

3.1.4 Deviance test (D) is goodness of fit test in MLR in equation 4.

$$D = 2\sum_{i=1}^{n}\sum_{j=1}^{j} O_{ij} \ln\left(\frac{O_{ij}}{E_{ij}}\right) \quad (4)$$

where O_{ij} is observed values
E_{ij} is expected values

3.1.5 The simplest optimizing method of discrimination was to maximize to posterior of correct allocation.

To obtain the posterior probability logit coefficients, the following equation is applied:

$$\log(\frac{\pi_i}{1-\pi_i}) = \beta_0 + \sum_{i=1}^{n} \beta_i x_i \quad (5)$$

Where β_0 is the intercept, β_i denotes the unknown logistic regression coefficients of x_i parameters and π_i denotes the probability that characteristic will occur. The quantity on the left side of Equation (5) is called a logit. The model can be generalized in the case where the dependent variables, unlike a binary logistic regression model, have more than two categories. Having '4' (stage IV) as the reference category, we can suppose c as the dependent variable with four categories, and the probability of being in category c (c='0' [Benign stage], c='1' [Stage I], c='2' [Stage II] and c='3' [Stage III]) is denoted by P(c) with the chosen reference category, P(4). For such a simple model, MLR with logit link can be represented as

$$\log\left(\frac{P(c)}{P(4)}\right) = \beta_0(c) + \sum_{i=0}^{3} \beta_i(c) x_i, c = 0..3 \quad (6)$$

In this model, the same independent variable appears in each of the c categories, and the separate intercept, $\beta_0(c)$, and slopes (or logit coefficients), $\beta_i(c)$ are usually estimated for selected parameters in each contrast. A way to interpret the effect of independent variables, x_i on the probability of being in category c, is to use predicted probabilities, P(c), for different values of x_i:

$$P(c) = \frac{\exp(\beta_0(c) + \sum_{1}^{n} \beta_i(c) x_i}{1 + \sum_{k=1}^{4} \exp\left(\beta_0(k) + \sum_{i=1}^{n} \beta_i(k) x_i\right)} \quad (7)$$

Then, the probability of being in the reference category, '4' (stage IV), can be calculated by subtraction:

$$P(4) = 1 - \sum_{k=0}^{3} P(k) \quad (8)$$

The category with the highest probability is the final prediction. For detailed descriptions on models with categorical data we refer to (Hayatshahi, S.H.S., 2005).

3.2. Classification

We wish to classify a patient into one specific class (for example, survival). For many purposes, it will be more helpful to know the predicted probability of survival. A simple but much neglected method is logistic regression which is specified by:

$$P(class2 \mid x) = \frac{e^{\lambda}}{1+e^{\lambda}} \quad (9)$$

Where $\lambda = \beta_0 + \beta_1 x_1 + ... + \beta_p x_p$

$$P(Class1 \mid x) = 1 - P(Class2 \mid x) = \frac{1}{1+e^{\lambda}} \quad (10)$$

$$\frac{P(class2 \mid x)}{P(Class1 \mid x)} = e^{\lambda} \quad (11)$$

The explanatory variables linearly control the log-odds λ in favour of class 2 (survival). The parameters β are chosen by maximum likelihood that is by maximizing the log-likelihood

$$L = \sum_i \log p(class_i \mid x_i) \quad (12)$$

By comparing the patients with features x and the feature patients, we will be able to predict P (class 2 | x), probability of survival.

Maximum likelihood is known as 'entropy' fitting and is definitely not common (and supported by amazingly few packages). It is more common to use the regression methods we discuss in section 2, which may be adequate for predicting the class (survival or death) but will be less good for predicting probabilities.

The extension to $k > 2$ classes is even less well known, although it has a long history. The idea is to take the log-odds of each class relative to one class, so the model becomes

$$\frac{P(Classj \mid x)}{P(Class1 \mid x)} = e^{\lambda j} \quad , \quad j = 1,2,...,k \qquad (13)$$

and so

$$P(classj \mid x) = \frac{e^{\lambda j}}{\sum_c^k e^{\lambda j}} \qquad (14)$$

With $\lambda_j = \beta_j^T x$ this is known as MLR. The parameters (β_j) are fitted by maximizing the log-likelihood L given in equal (2). There have been surprisingly few non-linear extensions in the statistics literature.

$$P(classj \mid x) = \frac{e^{\lambda j}}{1 + e^{\lambda j}} \quad , \; j = 1,...,k \qquad (15)$$

This is an appropriate model for diagnosis where a patient might have none, one or more out of k diseases, but not for general classification problems.

For the evaluation of the right classification in each model, two indices are: sensitivity and specificity. Sensitivity is defined as the capacity of an assessment instrument or battery to yield a positive result for a person with the diagnostic condition or attribute of interest. When using a medical analogy, sensitivity is the proportion of "diseased" individuals who obtain scores above the cut-off point of a screening test. That is:

$$Sensitivity = \frac{Diseased\ persons\ with\ positive\ test\ results}{All\ diseased\ persons}$$

$$\text{Sensitivity} = \text{TP} / (\text{TP}+\text{FN}) \qquad (16)$$

Similarly, specificity reflects the capacity of an assessment instrument to yield a negative result for a person without a diagnostic condition or attribute. That is, the proportion of "nondiseased" persons who obtain normal-range scores on the screening test equals specificity:

$$Specificity = \frac{Nondiseased\ persons\ with\ negative\ test\ results}{All\ non\ diseased\ persons}$$

$$\text{Specificity} = \text{TN} / (\text{TN}+\text{FP}) \qquad (17)$$

Both sensitivity and specificity were defined first by (Yerushalmy, 1947) and have been an important part of the medical literature since that time (Jobson, 1992).

Related calculations such as False positive rate, False negative rate, Likelihood ratio positive, Likelihood ratio negative, Accuracy, Positive predictive value (PPV), and negative prediction value (NPV) were done. The related calculation done is shown in Eq.18-23.

$$\text{False positive rate } (\alpha) = 1 - \text{specificity} = \text{FP} / (\text{FP} + \text{TN}) \qquad (18)$$

$$\text{False negative rate } (\beta) = 1 - \text{sensitivity} = \text{FN} / (\text{TP} + \text{FN}) \qquad (19)$$

$$\text{Likelihood ratio positive} = \text{sensitivity} / (1 - \text{specificity}) \qquad (20)$$

$$\text{Likelihood ratio negative} = (1 - \text{sensitivity}) / \text{specificity} \qquad (21)$$

$$\text{Positive predictive value (PPV)} = \text{TP}/(\text{TP}+\text{FP}) \qquad (22)$$

$$\text{Negative predictive value(NPV)} = \text{TN}/(\text{TN}+\text{FN}) \qquad (23)$$

where TP is true positive value, TN is true negative value, FP is false positive value and FN is false negative value.

5. Results

The data for this study were collected from May to September 2008. Data was collected at the Lopburi hospital in Thailand of 680 women.

5.1 Multinomial logistic regression

For experiments, the nine characteristics of breast cancer (Clump Thickness, Uniformity of Cell Size, Uniformity of Cell Shape, Marginal Adhesion, Single Epithelial Cell Size, Bare Nuclei, Bland Chromatin, Normal Nucleoli and Mitoses) for input values (x_1-x_9) which each characteristic containing number from 1-10 and five stages of breast cancer.

5.2 The statistical test in MLR:

Frequency and percentage distributions of the groups obtained from cluster analysis are presented in Table 1 and descriptive statistics of related independent variables are presented in Table 2.

Table 1 Percentage-frequency distribution of dependent variable in groups

Group	Frequency	%
Benign Stage	175	25.7
Stage I	29	4.3
Stage II	112	16.5
Stage III	36	5.3
Stage IV	328	48.2
Total	680	100

As Table 1 is evaluated, it is seen that the data set, which initially appeared homogeneous, actually consisted of five sub-groups with benign stage = 175(25.7%) patients in the first group, stage I = 29 (4.3%) in the second, stage II = 112 (16.5%) in the third group, stage III = 36(5.3%) in the forth, and stage IV = 328 (48.2%) in the fifth group.

Table 2 Descriptive statistics of Independent variable

Independent variable	$\bar{x}$	SD.
X_1	5.01	3.118
X_2	3.71	3.148
X_3	3.82	3.059
X_4	3.20	3.024
X_5	3.79	2.565
X_6	4.45	3.929
X_7	3.96	2.177
X_8	3.59	3.293
X_9	1.92	2.098
Overall Average	3.7166	2.9345

As Table 2 is evaluated, it is seen that x_1 had an average of 5.01±3.118; x_2 performed a different manner structure than x_1 and had an average of 3.71±3.148; x_3 had an average of 3.82±3.059; x_4 had an average of 3.20±3.024; x_5 had an average of 3.79±2.565; x_6 had an average of 4.45±3.929; x_7 had an average of 3.96±2.177; x_8 had an average of 3.59±3.293; and x_9 had an average of 1.92±2.098. Chi – square test of the hypotheses are as follows: Chi-square

Table 3 Goodness of fit test

	Chi-Square	df	p-value
Pearson	950.931	1024	0.949
Deviance	371.916	1024	1.000

Table 3 shows goodness of fit test and distribution test of data calculated by Chi-square test, p-value equal to 0.949 < 0.05. This model accepts H_0 Hypothesis meaning nine characteristics are independent variables which relate to stage of cancer.

Relationship between independent value and dependence value with logit function in Table 4.

Table 4 Model fitting information

Model	Model Fitting Criteria	Likelihood Ratio Tests		
	-2 Log Likelihood	Chi-Square	df	p-value
Intercept Only	1360.488			
Final	571.876	788.612	36	0.000

Table 4 shows model fitting information consisting of only constant values which yields -2LL = 1360.488, constant variables and independent variables which yields -2LL = 571.876. The final model is more suitable than intercept only model. (Chi-square = 1360.488-517.876 = 788.612 and p-value = 0.000 and p-value < 0.05). We reject H_0 Hypothesis which means the independent variables are related with the dependent variables.

5.3 Likelihood Ratio Test (LRT)

The likelihood ratio test (LRT) statistic is the ratio of the likelihood at the hypothesized parameter values to the likelihood of the data at the MLE.

Table 5 Likelihood ratio test

Effect	Model Fitting Criteria	Likelihood Ratio Tests		
	-2log Likelihood of Reduced Model	X^2	df	p-value
Intercept	819.992	248.115	4	.000
Clump Thickness (x_1)	608.852	36.976	4	.000
Uniformity of cell size (x_2)	573.752	1.876	4	.759
Uniformity of cell shape (x_3)	578.574	6.698	4	.153
Marginal Adhesion (x_4)	579.450	7.573	4	.109
Single Epithelial cell size (x_5)	573.045	1.168	4	.883
Bare nuclei (x_6)	615.165	43.289	4	.000
Bland Chromatin(x_7)	574.179	2.303	4	.680
Normal nucleoli (x_8)	575.491	3.614	4	.461
Mitoses (x_9)	580.243	8.367	4	.079

The Chi-square statistics is the difference in -2 log likelihoods between the final model and a reduced model. The reduced model is formed by omitting an effect from the final model. The null hypothesis is that all parameters of that effect are 0.

Table 5 shows coefficient of regression of model are x_2,x_3,x_4,x_5,x_7 and x_8 have significant value more than 0.05. MLR shows coefficient of model are x_1 and x_6 have significant less than 0.05 so that the prediction from log-likelihood function for classifying staging of breast cancer with P(Y≤4) of stage IV is the reference category. The model can be reduced as:

$$[P(Y=i)] = \ln\left(\frac{P_i}{1-P_i}\right)$$
$$= 819.992 + 608.852x_1 + 615.165x_6 \quad (24)$$

Testing between coefficient value and log – likelihood function in table 6 - 9.

Table 6 The significant parameters of Benign stage

Y=0	*β*	*Std Error*	*Wald*	*df*	*p-value*
Intercept	8.031	0.923	75.651	1	0.000
X_1	-0.616	0.123	25.020	1	0.000
X_2	0.118	0.174	0.461	1	0.497
X_3	-0.406	0.180	5.105	1	0.024
X_4	-0.260	0.123	4.477	1	0.034
X_5	-0.119	0.129	0.849	1	0.357
X_6	-0.405	0.080	25.738	1	0.000
X_7	-0.182	0.154	1.404	1	0.236
X_8	-0.067	0.094	0.511	1	0.475
X_9	-0.449	0.255	3.103	1	0.078

Table 6 shows the small p-values obtained for x_1, x_3, x_4 and x_5 indicating that they are the most significant predictor of benign stage in the model and the remaining parameters including size and shape of tumors as well as associated features not significant at level of 0.05.

Table 7 The significant parameters of stage I

Y=1	*β*	*Std Error*	*Wald*	*df*	*p-value*
Intercept	21.434	1.021	441.028	1	0.000
X_1	-0.589	0.162	13.235	1	0.000
X_2	0.312	0.339	0.847	1	0.357
X_3	-0.562	0.317	3.147	1	0.076
X_4	-0.158	0.203	0.606	1	0.436
X_5	-0.087	0.197	0.196	1	0.658
X_6	-0.464	0.157	8.707	1	0.003
X_7	-0.112	0.218	0.265	1	0.606
X_8	-0.230	0.201	1.305	1	0.253
X_9	-15.812	0.000	0.000	1	0.000

Table 7 shows the small p-values obtained for x_1, x_6 and x_9 indicating that they are most significant predictor of stage I of breast cancer in the model and the remaining parameters including size and shape of tumors as well as associated features not significant at level of 0.05.

Table 8 The significant parameters of stage II

Y=2	*β*	*Std Error*	*Wald*	*df*	*p-value*
Intercept	7.574	0.941	64.734	1	0.000
X_1	-0.646	0.127	25.671	1	0.000
X_2	0.193	0.192	1.008	1	0.315
X_3	-0.329	0.191	2.967	1	0.085
X_4	-0.337	0.146	5.307	1	0.021
X_5	-0.108	0.139	0.610	1	0.435
X_6	-0.432	0.090	22.748	1	0.000
X_7	-0.178	0.162	1.213	1	0.271
X_8	-0.119	0.109	1.189	1	0.276
X_9	-0.402	0.262	2.343	1	0.126

Table 8 shows the small p-values obtained for x_1, x_4 and x_6 indicating that they are most significant predictor of stage II of breast cancer in the model and the remained parameters including size and shape tumor as well as associated features not significant at level of 0.05.

Table 9 The significant parameters of stage III

Y=3	*β*	*Std Error*	*Wald*	*df*	*p-value*
Intercept	6.607	1.086	36.979	1	0.000
X_1	-0.550	0.152	13.084	1	0.000
X_2	-0.093	0.307	0.091	1	0.763
X_3	-0.562	0.295	3.621	1	0.057
X_4	-0.186	0.198	0.883	1	0.347
X_5	-0.176	0.202	0.764	1	0.382
X_6	-0.278	0.118	5.597	1	0.018
X_7	-0.295	0.217	1.851	1	0.174
X_8	0.103	0.143	0.518	1	0.472
X_9	-0.434	0.360	1.454	1	0.228

Table 9 shows the small p-values obtained for x_1 and x_6 indicating that they are the most significant predictors of stage III of breast cancer in the model and the remaining parameters including size and shape of tumors as well as associated features not significant at level of 0.05.

Tables 6-9 indicate statistical results for the significant parameters of logit stage 0 to stage 3 of breast cancer in MLR, receiving parameters in Classification Model as follows :

$$\ln\left[\frac{P_0}{1-P_0}\right] = 8.031 - 0.616x_1 - 0.46x_3 - 0.26x_4 - 0.405x_6 \quad (25)$$

From Eq. (25), it is found that clump thickness (x1), uniformity of cell shape (x3), marginal adhesion (x4) and bare nuclei (x6) prognosis of breast cancer patients were not present. Breast cancer may be up to stage 4.

$$\ln\left[\frac{P_1}{1-P_1}\right] = 21.434 - 0.589x_1 - 0.464x_6 - 15.812x_9 \quad (26)$$

From Eq. (26), it is found that clump thickness (x1), bare nuclei (x6) and Mitoses (x9) can predict breast cancer patients at stage 1 and maybe up to stage 4.

$$\ln\left[\frac{P_2}{1-P_2}\right] = 7.574 - 0.464x_1 - 0.337x_4 - 0.432x_6 \quad (27)$$

From Eq.(27), it is found that clump thickness (x1), marginal adhesion (x4) and bare nuclei (x6) can predict breast cancer patients at stage 2 and may be up to stage 4.

$$\ln\left[\frac{P_3}{1-P_3}\right] = 6.607 - 0.55x_1 - 0.278x_6 \quad (28)$$

From Eq.(28), it is found that clump thickness (x1) and bare nuclei (x6) can predict breast cancer patients at stage 3 and may be up to stage 4. We can conclude that clump thickness (x1) and bare nuclei (x6) are key factors that cause breast cancer in stage 4.

From Eq.(25) - Eq.(28), it shows Beta values at the p-value < 0.05, which is considered a condition affecting the stage of cancer.

5.4 Classification of MLR

Table 10 Classification

Observed	**Benign**	**Stage 1**	**Stage 2**	**Stage 3**	**Stage 4**	**%**
0	157	2	13	3	0	89.7
1	26	2	1	0	0	6.9
2	98	0	14	0	0	12.5
3	33	0	0	3	0	8.3
4	0	0	0	0	328	100
Overall Percentage	**46.2**	**0.6**	**4.1**	**0.9**	**48.2**	**74.1**

From Table 10, the correct classification for the benign stage of breast cancer is 46.2% and stage 4 is 48.2%. The overall percentage is 74.1%.

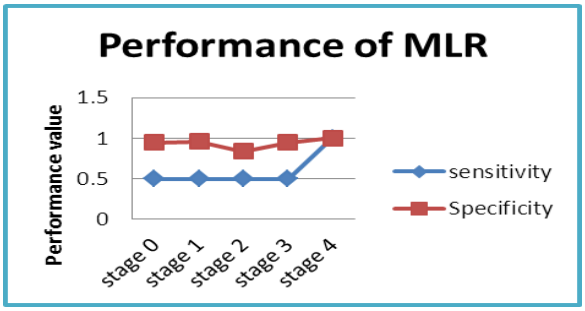

Figure 1 Performance classification of MLR

5.5 Performance Classification of MLR

Two good indices for the evaluation of the right classification by each model are sensitivity and specificity.

Figure 1 shows the sensitivity of close to 1, which shows that there has been the prediction accuracy in all stages, but it is noted that specificity value in stage 2 may affect treatment response which was not due to cancer; stage 2 is expected to be pathogenic which is harmful to the patient.

Related calculations such as False positive rate, False negative rate, Likelihood ratio positive, Likelihood ratio negative, Accuracy, Positive predictive value (PPV) and Negative prediction value (NPV) were done. The following calculations done are shown in table 11 (see below).

Table 11 Related calculation of Performance

Stage	Sensitivity	Specificity	False positive rate	False negative rate	Likelihood ratio positive	Likelihood ratio negative	Accuracy	PPV	NPV
Benign	0.50	0.95	0.05(5%)	0.05(5%)	10	0.526	0.742	0.897	0.689
Stage 1	0.50	0.96	0.04(4%)	0.05(5%)	12.5	0.520	0.957	0.068	0.996
Stage 2	0.50	0.84	0.15(15%)	0.05(5%)	3.31	0.588	0.835	0.125	0.975
Stage 3	0.50	0.95	0.04(4%)	0.05(5%)	10.2	0.525	0.947	0.083	0.995
Stage 4	1.00	1.00	0(0%)	0(0%)	-	0	1.000	1.000	1.000

6. Conclusion

The results obtained are in accordance with actual survival rates. The classification of stages of breast cancer is consistent with data used in 680 cases. This study proposes MLR to predict five stages of breast cancer (Benign, I, II, III and IV) and the nine characteristics of breast cancer: Clump Thickness (X_1); Uniformity of Cell Size(X_2); Uniformity of Cell Shape (X_3); Marginal Adhesion (X_4); Single Epitheliai Cell Size (X_5); Bare Nuclei (X_6); Bland Chromatin (X_7); Normal Nucleoli (X_8); and Mitoses (X_9) are used as independent variable. Based on the results, Multinomial Logistic Regression (MLR) shows x_1 and x_6, which is a coefficient of a model with significance less than 0.05. As such, the prediction of log – likelihood function for classification staging of breast cancer with P(Y≤4) of stage IV is a reference category. It is thus reduced as a model as:

$$\log\left(\frac{P_i}{1-P_i}\right) = 819.992 + 608.852x_1 + 615.165x_6$$

The performance of classification in this model is measured, and this reveals sensitivity and specificity. The results are specificity of 0.9428 and sensitivity of 0.9537. The correct classification for the benign stage of breast cancer is 46.2%, and the stage 4 is 48.2%. The overall percentage is 74.1%. The result relative of performance is shown in table 11.

This study is beneficial to the treatment of breast cancer because the method used to analyze the predictive accuracy of Benign stage cancers was 89.7%, resulting in increasing survival rates of patients with breast cancer. From equation 25-28, it is notable that the X9 (Mitoses) is an indication of a patient with cancer in stage 2, so this parameter gives value to medical treatment, leading to feasible treatment for patients with abrupt breast cancer.

7. Acknowledgements

The authors would like to express gratitude to Associate Professor Dr. Chom Kimpan for the permission for the extension of this research.

8. References

Agreti, A. (2002). Categorical Data Analysis, (2nd ed.). New York, USA: *John Wiley & Sons.*

Bandhita, P. & Noparat, T. (2006, June). Ordinal Regression Analysis in factors related to Sensorial Hearing Loss of the Employee. *Industrial factory in Lampang Thailand. Mathematic, Statistics and Their Application, Penang.*

Hayatshahi, S. H. S., Abdolmaleki, A., Safarian, S., & Khajeh, K. (2005, October). Non-linear quantitative structure-activity relationship for adenine derivatives as competitive inhibitors of adenosine deaminase. *Biochem Biophys Research Comunication., 338*, 1137-1142. Retrieved from http://elsevier.com/locate/ybbrc.

Hosmer, D. W. & Lemeshow, S.(2000). Application of logistic regression. New York, USA: *John Wiley & Sons.*

Jobson , J. D. (1992). Applied multivariate data analysis, (2nd ed.). New York, USA/ Berlin Heidelberg: *Springer*.

Long, J. S. (1997). Regression Models for categorical and Limited Dependent Variables, USA: *Sage Publication, Inc.*

Stephenson, B. (2008). Chapter 3 Binary response and logistic regression analysis. Retrieved from http:// public.iastate.edu/~stat415/ *stephenson/*

Swets, J. (1988). Measuring the accuracy of diagnostic systems. *Science, 240,* 1285-1293. doi : 101126/science.3287615

Swets, J. A., Dawes, R. M. & Monahan, J. (2000, October). Better decisions through science. *Scientific American, 283*, 82-87.

Vecchio, T (1966). Predictive value of a single diagnostic test in unselected populations. *New England Journal of Medicine, 274,* 1171-1173.

Wingo, P.A., Tong, T. & Bolden, S. (1995). Cancer statistics, *CA Cancer J Clin, 46*(1), 5-27. doi : 10.3322/canjclin.46.1.5