European Journal of Molecular & Clinical Medicine
ISSN 2515-8260 Volume 7, Issue 8, 2020

ANALYSIS OF PROGNOSIS FACTORS OF BREAST CANCER PATIENTS THRU SOCIO ECONOMIC PROFILE: AN APPLICATION OF BINARY LOGISTIC REGRESSION

Mrinal Deka
Ph.D. Research Scholar, Department of Statistics, Assam University, Silchar, India

Dibyojyoti Bhattacharjee
Professor, Department of Statistics, Assam University, Silchar, India

Abstract
Background: The idea of Prognosis factor is based on the variables that can be used to assess the chance of recovery from a disease. It may also be defined as the prior knowledge about any disease before treatment.

Method: In this paper, selective prognostic factors (Age, Node and Tumour size) are analysed by logistic regression in patients who are suffering from Breast cancer based on data collected from the Cachar Cancer Hospital and Research Centre, Silchar, Assam, India. The purpose of the research is to analyse the effect of the prognosis factors on the remission of breast cancer; separately for economically weaker as well as well to do patients.

Results: The study claims that there are 50.1 percent and 65.8 percent chance of remission of cancer for patients of age above 50 in case of breast cancer with below the poverty line and above the poverty line respectively.

The present study has considered the cutoff value of 2 cm as the determining prognostic factor in relation to tumour size. Thus, the chance of remission from cancer is 22.3 percent and 34.7 percent for below poverty line and above poverty line patients, respectively, if tumour size is greater than 2 cm.

It also endeavours to ascertain that there are 10.9 percent and 18.1 percent chance of remission of cancer, if the disease has metastasized to regional lymph nodes, for below poverty line and above poverty line patients respectively.

Conclusion: There is a significant difference between the two poverty lines (APL and BPL) in terms of node and tumour size of breast cancer. The increasing sizes of tumour and node have lesser chance to follow-up as well as poorer survival and has a significant difference for patients belonging to both the poverty levels. The prognosis factors have the significant impact on the remission of breast cancer and depends on the socio-economic status of the patients due to the different standard of living, tendency of early diagnosis and the awareness level of cancer disease.

Keywords: Prognosis Factor; Logistic Regression; Breast Cancer; Bio Statistics.

INTRODUCTION

Prognosis factors are defined as the variables that can be used to assess the chance of recovery from a disease. It is also considered as the prior knowledge about any disease before treatment. The concept is applied to the cancer patients to get an idea of how cancer will affect the body and how it will respond to the therapies. It is difficult for the common people to take decisions about treatment methods due to lack of knowledge and different socio-economic backgrounds. Thus, the prior knowledge will benefit the common people to participate in clinical and health policy decisions through prognostic and economic evaluation of cancer treatments. Although many prognostic factors have been identified over the last few decades, affecting the survival outcomes for breast cancer, there are some that have been proven to be of definite significance through various statistical methods. These include – tumour size, nodal status, distant metastasis, histologic grade, histologic type, mitotic figure counts, hormone receptor status like Estrogen Receptor (ER) and Progesterone Receptor (PR) positivity and age of the patient *(Russo.et.al, 1987).* The present study will focus on tumour size, nodal status and age in relation to survival outcomes for breast cancer. Though these factors are already proven to be of significant importance in various studies, the present study will bring out their importance so as to guide future treatment decisions. In addition, the economic status of the patients has been evaluated as a possible demographic factor affecting choice of treatment and the resulting survival outcomes.

It is a matter of serious concern that new cases of the breast cancer are growing up day by day all over the world *(Ali.et.al, 2011)*. Thus, the effect of the prognosis factors on cancer should be studied scientifically so that the patients can get an idea of economic and health policy decisions during the ongoing treatment processes. A comparative study was performed by *Kroman.et.al,* in the year 2000, among the treated vs. not treated through adjuvant cytotoxic treatment in case of breast cancer. The study revealed that there is a negative prognostic effect of young age in women diagnosed with low risk disease who did not receive adjuvant cytotoxic treatment (*Kroman.et.al,* 2000). Another study estimated the variations of Out-of-Field Dose, that are associated with radiotherapy, for the different parameters like field size and depth of cancer using the Markus Ionization Chamber Detector *(Abdelaal.et.al, 2020).* Again, in the year 2018, a study concluded that the effects of blood pooling have the impact on the levels of radioactivity measured in cancer tissues *(Yavari K, 2018).* A research was conducted in the recent year and concluded that Saliva officinalis can potentially prevent breast cancer *(Zare H,2019).*

Economic burden plays an important role in the growing incidence of cancer as the cost of treatment is expensive and the income losses are significant *(Nair.et.al, 2014).* It seems logical that economically weaker patients are not in the favour of early diagnosis of cancer due to the high cost of the treatments, thus results in poor chance of survival. Taking treatment at early stage of cancer can relief the pain as well prevent the cancer from metastasis *(Sun.et.al, 2017).* The patients are not often informed about the cause, nature and cost of therapies/treatments. Similarly, patients mostly do not have any information on the nature of prognostic factors and this lack of information may lead to the more advanced stages of cancer at diagnosis *(Caplan L, 2014).* The 'Stage' of the cancer is an important risk factor as due to longer time between the onset of cancer symptoms and the patient's presentation to health care, leading to later-stage diagnoses and therefore less eligibility for potentially curative treatment *(Walter.et.al, 2015).*

One study claims that the efforts to promote early detection should be continued in fighting with breast cancer as the primary prevention of breast cancer is still not available *(Caplan L, 2014).* Therefore, serious research is needed on the prognosis factors of breast cancer, which might help the patient/family members and the clinics/hospitals to take optimal decisions for cancer treatments. Although, different researchers have successfully brought up the importance of prognosis factors for different sites of cancer, but there is still a lack of studies on the relationship between the prognosis factors and the socio-economic status of the patients. Thus, keeping these points in mind, the present research is structured to get a conclusion of how the prognosis factors effect on the chance of remission of breast cancer; separately for economically weaker as well as economically well-off patients. The findings of the present research work will benefit the society for better treatment of cancer.

OBJECTIVE OF THE STUDY

The paper is designed keeping in view the following objective:

- To study the effect of prognosis factors on the chance of remission of breast cancer; separately for economically weaker and economically well-off patients.

METHODS

The data used for the present study is secondary in nature collected from the Cachar Cancer Hospital and Research Centre, Silchar, Assam, India for the year 2013-2019. The dataset is stratified on the basis of the economic condition of the patients (BPL – below poverty line and APL – above poverty line). The classification is as the patients included in this study belong to different economic backgrounds. Thus, the patients, who have the BPL card[1] are considered as Below Poverty Line category and those without the card are considered to be an Above Poverty Line category. The following table gives an overview of the dataset.

Table 1: Classification of the patients considered in the study on the basis of clinical and demographic parameters

Sites of Cancer	Variable		Frequency	Percentage
Breast Cancer (Below Poverty Line)	**Gender**	**Male**	**0**	**0 %**
		Female	**200**	**100 %**
	Age	**Age above 50**	**113**	**56.5%**
		Age below 50	**87**	**43.5%**
	Lymph Node	**Cancer spreads to LN**	**70**	**35%**
		Cancer not spreads to LN	**130**	**65%**
	Tumour size	**Tumour greater than 2 cm**	**163**	**81.5%**
		Tumour less than 2 cm	**37**	**18.5%**
Breast Cancer (Above Poverty Line)	**Gender**	**Male**	**0**	**0%**
		Female	**200**	**100%**
	Age	**Age above 50**	**93**	**46.5%**
		Age below 50	**107**	**53.5%**
	Lymph Node	**Cancer spreads to LN**	**95**	**47.5%**
		Cancer not spreads to LN	**105**	**52.5%**
	Tumour size	**Tumour greater than 2 cm**	**146**	**73%**
		Tumour less than 2 cm	**54**	**27%**

In this study, data were collected from a total of 400 female patients, half of which belongs to BPL group and the other half to APL group. There seems to be the difference between the two poverty lines in terms of size of the tumour as well as node status and thus z test for two sample proportion has been performed. It is appropriate for the comparison between the groups in proportion *(Kleinbaum and Kupper, 1978)* and the test statistics can be computed as,

$$\hat{z}_{node\_status} = \frac{(\hat{p}_1 - \hat{p}_2) - 0}{\sqrt{\hat{p}_{node\_status}(1 - \hat{p}_{node\_status})\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}} \quad ...(1)$$

$$\hat{z}_{tumour\_size} = \frac{(\hat{p}_3 - \hat{p}_4) - 0}{\sqrt{\hat{p}_{tumour\_size}(1 - \hat{p}_{tumour\_size})\left(\frac{1}{n_3} + \frac{1}{n_4}\right)}} \quad ...(2)$$

where, $\hat{p}_1 =$ proportion of node status for APL group

$\hat{p}_2 =$ proportion of node status for BPL group

$\hat{p}_{node\_status} =$ overall proportion of node status for both group (APL and BPL)

[1] BPL or Ration card is an official card issued by the government of India for economically weak households to support in terms of food, fuel or other essential goods.

n_1 = sample size for APL group in terms of node status

n_2 = sample size for BPL group in terms of node status

$\hat{p}_3$ = proportion of tumour size for BPL group

$\hat{p}_4$ = proportion of tumour size for APL group

$\hat{p}_{tumour\_size}$ = overall proportion of tumour size for both group (APL and BPL)

n_3 = sample size for BPL group in terms of tumour size

n_4 = sample size for APL group in terms of tumour size

The null hypothesis for the test statistic can considered as,

H_0 : the proportions are same

against the alternative hypothesis

H_1 : the proportions are not the same.

Again, the difference between the two poverty lines (APL and BPL) has been tested in terms of the sizes of the tumour, the node and the follow-up period of the breast cancer patients. We are interested to observe the progressing sizes of nodal and tumour have any impact on the follow-up period of the breast cancer patients in both the groups (APL and BPL) separately. Thus, the multiple regression technique is applied and the models can be written as,

$$Follow\_up\_period_{APL} = \beta_0 + \beta_1 \times Node\_size + \beta_2 \times Tumour\_size ...(3)$$

$$Follow\_up\_period_{BPL} = \beta_0 + \beta_1 \times Node\_size + \beta_2 \times Tumour\_size ...(4)$$

where $\beta_0's$ are the constants and $\beta_1's$, $\beta_2's$ are the regression coefficients of Node and Tumour sizes for APL and BPL group respectively.

Before including the independent variables (Node size and Tumour size) in the regression models, one important point should keep in mind that some independent variables that can be included in the models may play a redundant role, which could direct effect on the models and thus cannot be considered as reliable. Thus, Variance Inflation Factor (VIF) technique is applied through Multicolinearity Analysis.

The z test statistic is also computed to test the statistical difference between the regression equations in terms of node and tumour size and it should be noted the 'test of normality' must be performed before the z test *(Kleinbaum and Kupper, 1978).* The statistics can be computed as,

$$\hat{z}_{node\_status} = \frac{b_1 - b_2}{S_{b_1 - b_2}} ...(5)$$

$$\hat{z}_{tumour\_size} = \frac{b_3 - b_4}{S_{b_3 - b_4}} ...(6)$$

where, b_1 = regression coefficient associated with node size of APL group

b_2 = regression coefficient associated with node size of BPL group

$S_{b_1-b_2} = \sqrt{S_{b_1}^2 + S_{b_2}^2}$ = the difference of standard error between the regression coefficients associated with node size of APL group and BPL group.

$b_3 =$ regression coefficient associated with tumour size of APL group

$b_4 =$ regression coefficient associated with tumour size of BPL group

$S_{b_3-b_4} = \sqrt{S_{b_3}^2 + S_{b_4}^2}$ = the difference of standard error between the regression coefficients associated with tumour size of APL group and BPL group.

Now, Binary logistic regression model is applied to observe the probability of remission of disease in the presence of these prognostic factors (Age, Node and Tumour size). It is the chance of an event occurring in a model based on individual characteristics. Because the chance is mainly a ratio, it can be computed as:

$$p = \frac{e^{\beta_0+\beta_1 x_1+\beta_2 x_2+\ldots+\beta_m x_m}}{1+e^{\beta_0+\beta_1 x_1+\beta_2 x_2+\ldots+\beta_m x_m}} \ldots(7)$$

where p indicates the probability of an event, β_i's are the regression coefficients associated with the reference group and x_i's are the explanatory variables *(Francis R., 2017)*.

In the present study, it will be convenient if we write our binary logistic regression model as follows:

$$p = \frac{e^{\beta_0+\beta_1 x_1+\beta_2 x_2+\beta_3 x_3}}{1+e^{\beta_0+\beta_1 x_1+\beta_2 x_2+\beta_3 x_3}} \ldots(8)$$

where p = Probability of remission of cancer

$x_1 =$ Age of the patients

$x_2 =$ Node status of the cancer

$x_3 =$ Tumour size of the cancer

Thus, we can write the model as follows.

$$\log it(p) = \beta_0 + \beta_1 \times Age + \beta_2 \times Node\_status + \beta_3 \times Tumour\_size \ldots(9)$$

The following flowchart gives a better understanding of the Binary Logistic Regression model.

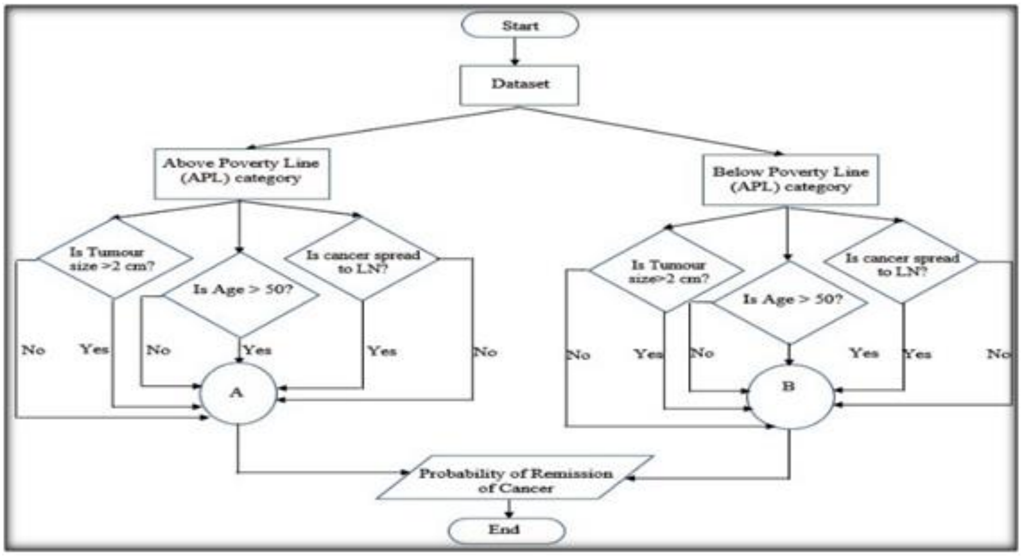

Fig. 1: Flowchart of the Binary Logistic Regression model

RESULTS

At the outset, z test for two sample proportion has been performed to test the difference between the two poverty lines (APL and BPL) in terms of tumour size and node status of breast cancer. The following table provides the results of z test for proportion of both the poverty lines separately.

Table 2: Result of z test for proportion of poverty lines in terms of Node status and Tumour size

Factors	Poverty line	Sample proportions	z test statistic	p value
Lymph Node (LN) (Cancer spreads to LN)	Above Poverty Line Below Poverty Line	0.475 0.350	2.5391	0.005
Tumour size (greater than 2 cm)	Above Poverty Line Below Poverty Line	0.730 0.815	2.0276	0.020

It is obtained from **Table 2** that there is a significant difference between the two poverty lines (APL and BPL) in terms of node status and tumour size of breast cancer as the p value is less than 0.05.

Now, Variance Inflation Factor (VIF) technique is applied through Multicolinearity Analysis for redundancy test among the independent variables. VIF technique assesses whether any independent variables are redundant in the model, which could direct effect on the model and thus cannot be considered as reliable. The following table gives the idea of redundancy among the independent variables.

Table 3: VIF values for the independent variables in the regression model

Site of Cancer	Poverty Lines	Variables	VIF values
Breast Cancer	Above poverty Line	Node size	1.056
		Tumour size	1.031
	Below poverty Line	Node size	1.087
		Tumour size	1.092

We can conclude from the **Table 3** that the independent variables (Node size and Tumour size) are not redundant and can be included in the regression model as VIF values less than 5.

Thus, the multiple regression technique is applied to observe whether the progressing sizes of a node and tumour have any impact on the follow-up period of the breast cancer patients in both the groups (APL and BPL) separately. The following table provides the result of multiple regression model.

Table 4: Result of Multiple Regression Model

Poverty line	Factors	Intercept	Regression coefficient	Standard error	p value
Above Poverty Line (APL)	Node size Tumour size	6.516	-0.169 -0.067	0.009 0.003	0.000 0.000
Below Poverty Line (BPL)	Node size Tumour size	5.516	-0.081 -0.019	0.013 0.024	0.000 0.000

From **Table 4**, we observe that the progressing sizes of a node and tumour have the significant impact on the follow-up period of the breast cancer patients in both the groups (APL and BPL) as the p values are less than 0.05. In other words, we can state that the patients with increasing sizes of tumour and node have less chance to follow-up as well as poor survival for both the poverty line (APL and BPL).

We compute z test statistic to test the significant difference between the regression equations in terms of node and tumour size and it should be noted the 'test of normality' must be performed before the z test. The following table provides the result of test of normality.

Table 5: Test of Normality

Factors	Poverty line	Kolmogorov-Smirnov test	p value	Shapiro-Wilk test	p value
Node size	Above Poverty Line Below Poverty Line	0.248 0.217	0.580 0.587	0.795 0.842	0.365 0.387
Tumour size	Above Poverty Line Below Poverty Line	0.246 0.251	0.389 0.307	0.792 0.794	0.613 0.651

It is assumed the null hypothesis that both the groups (APL & BPL) follow a normal distribution in terms of node status and tumour size of breast cancer and the results show that the assumption of normality is valid in each case as p values are greater than 0.05.

Table 6: Results of comparison between the regression coefficients

Factors	Poverty line	Regression coefficient	Standard error	z test statistic	p value
Node size	Below Poverty Line Above Poverty Line	-0.081 -0.169	0.013 0.009	5.5696	0.000
Tumour size	Below Poverty Line Above Poverty Line	-0.019 -0.067	0.024 0.003	1.9846	0.025

The significant results have obtained from the above table and we can conclude that the regression coefficients differ significantly as p values are less than 0.05. In other words, we may conclude that there is a significant difference between the two poverty lines (APL and BPL) in terms of the sizes of the tumour, the node and the follow-up period of the breast cancer patients.

Now, logistic regression model is performed to observe the probability of remission from the cancer in the presence of the three prognosis factors viz. Age, Node and Tumour size. The analyses are done separately for both categories (BPL and APL) of breast cancer as both the groups differ significantly.

Table 7: Analysis of logistic regression model for breast cancer with below poverty line

Prognosis Factors	B	S.E	Wald	df	Sig.	Odds ratio	95 % CI	
							Lower	Upper
Age above 50 Yes	0.663	0.291	5.203	1	0.023	1.501	1.098	3.432

No®								
Cancer spread to Lymph Node Yes No®	0.737	0.333	4.897	1	0.027	1.109	1.088	4.016
Tumour size is greater than 2 cm Yes No®	0.889	0.331	7.208	1	0.007	1.223	1.210	4.654
Constant	0.392	0.259	2.293	1	.0130	1.480	-	-

Note: ® Denotes reference category.

We have obtained from **Table 7** that in case of breast cancer with below poverty line category, there is a 50.1 percent chance of remission from the cancer if the patient age is above 50, which is better in comparison to that in patient age below 50. Again, if the cancer has already metastasized to axillary lymph nodes, the chance of remission from the cancer is only 10.9 percent, which is worse as compared to the cancer that has not spread to the lymph nodes. The patients with tumour size greater than 2 cm have a 22.3 percent chance of remission from cancer after the treatment, which is worse in comparison to the patients with the size of tumours less than 2 cm.

Table 8: Analysis of logistic regression model for breast cancer with above poverty line

Variables	**B**	**S.E**	**Wald**	**df**	**Sig.**	**Odds ratio**	**95 % CI**	
							Lower	**Upper**
Age above 50 Yes No®	1.043	0.366	8.111	1	0.004	1.658	1.384	5.817
Cancer spread to Lymph Node Yes No®	1.051	0.516	4.152	1	0.042	1.181	1.041	5.864
Tumour size is greater than 2 cm Yes No®	1.299	0.555	5.484	1	0.019	1.347	1.236	3.877
Constant	-2.398	0.522	6.083	1	0.000	0.091	-	-

Note: ® Denotes reference category.

Similarly, from **Table 8**, it is obtained that there is 65.8 percent more chance of remission from the cancer if the patient age is above 50 in the comparison to the patient of age below 50. Again, if the cancer has already metastasized to axillary lymph nodes, the chance of remission from the cancer is only 18.1 percent as compared to the cancer that has not metastasized to regional lymph nodes. The patients with tumour size greater than 2 cm have a 34.7 percent chance of remission of cancer in the comparison to the patients with the size of tumours less than 2 cm in the case of breast cancer with above poverty line category.

DISCUSSION

Preclusion of cancer is one of the most significant public health challenges of the 21st century *(Ali et. al, 2011)*. Further, new cases of breast cancer show an ever increasing incidence all over the world. Several prognostic factors have been identified which affect the outcomes of disease and treatment. Thus, the present study is mainly focused on the chance of remission of cancer in relation to three prognostic factors viz. Age, Node and Tumour size for the patients that are suffering from breast cancer.

The age of the patient is a well-defined prognosis factor for local recurrence. It is well established that patient age greater than 35 or 40 is associated with an increased frequency of local recurrence due to presence of various adverse pathologic features, such as lymph vascular invasion, grade 3 histology, absence of Estrogen Receptor (ER) and Progesterone Receptor (PR), presence of HER2 and presence of extensive intra-ductal component *(Kollias.et.al, 1997).* Our study finds that there are 50.1 percent and 65.8 percent chance of remission of cancer for patients of age above 50 in case of breast cancer with below the poverty line and above the poverty line respectively. Thus, in agreement with previous studies, we can consider the higher age as a good prognostic factor and younger age as a poor prognostic factor in breast cancer.

Tumour size is considered to be the best measure of tumour behaviour in breast cancer. Patients with a primary tumour size of less than 1 cm exhibit a frequency of only 10 percent to 20 percent of nodal metastasis, such that the 10-year disease-free survival rate is about 90 percent *(Carter.et.al, 1989)*. Since the American Joint Committee on Cancer (AJCC) has described tumour size less than or equal to 2 cm as T1, the present study has focused on this cutoff value of 2 cm as the determining prognostic factor in relation to tumour size. The chance of remission from cancer is 22.3 percent and 34.7 percent for BPL and APL patients, respectively, if tumour size is greater than 2 cm.

Axillary lymph node status has been described as the second most important prognostic factor in relation to disease-free survival, as well as overall survival in breast cancer. 70 percent of node positive patients are likely to develop a recurrence compared to only 20 percent to 30 percent of node-negative patients *(Veronesi.et.al, 1993).* Patients with 4 or more numbers of involved nodes have a worse prognosis when compared to those with less than 4 nodes *(Fisher.et.al, 1993).* It is found from our study that there are 10.9 percent and 18.1 percent chance of remission of cancer, if the disease has metastasized to regional lymph nodes, with below the poverty line and above poverty line respectively. Thus, we conclude that the parameters viz. Tumour size and Node status, which measure the stage of the cancer patients is significant prognostic factors that help in predicting tumour behaviour and survival outcomes.

Looking at the global scenario and studying the cancer incidence related spatial data, economic conditions of the patients are also a matter of serious concern *(Nair.et.al, 2014).* Since the cost of treatment is expensive and the income losses are significant, studies on the socio-economic status of the patients is expected to add lots of value to the particular research. It seems logical that economically stronger patients have the tendency in terms of awareness level of cancer disease, early diagnosis, better quality of life and thus result in good prognosis as well as rapid remission from the disease. Previous studies also successfully brought up the significant relationship between the prognosis factors and the patients' standard of living, tendency of early diagnosis and the awareness level of cancer disease *(Agarwal.et.al, 2017 and Meneses.et.al, 2012).* Another study claims that patients, mostly do not have any information on the nature of prognostic factors and this lack of information may lead to the more advanced stages of cancer at diagnosis *(Caplan L, 2014).*

The present study also finds the significant difference between the poverty lines (APL and BPL) with respect to the selective prognosis factors. It claims that there are 50.1 percent, 22.3 percent and 10.9 percent chance of remission from breast cancer in the case of below poverty line for the prognosis factors of Age, Tumour and Node size respectively. The percentages are higher for the patients of above poverty line category with 65.8 percent, 34.7 percent and 18.1 percent chance of remission from breast cancer. Thus, in agreement with previous studies, we may conclude that prognosis factors have the significant impact on the remission of breast cancer and it varies from socio-economic status of the patients due to the different standard of living, tendency of early diagnosis and the awareness level of cancer disease as mentioned in the previous studies.

CONCLUSION

The present study has successfully brought out the relationship between specific prognostic factors and survival outcomes in breast cancer with respect to the different socio-economic status of the patients. There is a significant difference between the two poverty lines (APL and BPL) in terms of node and tumour size of breast cancer. The increasing sizes of tumour and node have lesser chance to follow-up as well as poorer survival and has a significant difference for both the poverty lines. The application of logistic regression model to the selected prognostic factors has revealed that a tumour size greater than 2 cm denotes a dismal prognosis, as does the presence of axillary nodal involvement. Also, patient's age less than 50 years is associated with a worse prognosis and poor overall survival. The prognosis factors have the significant impact on the remission of breast cancer and it vary from socio-economic status of the patients due to the different standard of living, tendency of early diagnosis and the awareness level of cancer disease.

We expect that the present study will pave ways for further study on the topics and would be beneficial for the researchers in the field. Further research would be challenging to study the health financing methods of different socio-economic groups and suggest policies related to tailor made insurance and expenditure management for the cancer patients. Additional challenge might be to decompose the cost component (like doctors' fees, therapy related cost, etc.) according to the incidence of cancer and study the pattern over time.

REFERENCES

1. Abdelaal M.A, Attalla M.E, Elshemey M.W. Estimation of Out-of-Field Dose Variation using Markus Ionization Chamber Detector. SciMedicine J. 2020. doi: http://dx.doi.org/10.28991/ SciMedJ-2020-0201-2
2. Agarwal S, Ying J, Boucher M.K. The association between socioeconomic factors and breast cancer-specific survival varies by race. Plos One. 2017. doi: https://doi. org/10.1371/journal.pone.0187018
3. Ali I, Wani A.W, Sallem K. Cancer scenario in India with future perspectives. Cancer Therapy.2011;8:50-70. Available from: http://licgreaternoida.com/data/documents/ CancerTherapy2011.pdf
4. Caplan L. Delay in breast cancer: implications for stage at diagnosis and survival. Front. Public Health. 2014. doi: https://doi.org/10.3389/fpubh.2014.00087
5. Carter CL, Allen C, Henson DE. Relation of tumour size, lymph node status and survival in 24740 breast cancer cases. Cancer. 1989;63:181–187. doi: 10.1002/1097-0142(19890101)63:1<181::aid-cncr2820630129>3.0.co;2-h.
6. Fisher ER, Anderson S, Redmond C, Fisher B. Pathologic findings from the national surgical adjuvant breast project protocol B-06: 10-year pathologic and clinical prognostic discriminants. Cancer. 1993;71:2507–2514. doi: 10.1002/1097-0142(19930415)71:8<2507::aid-cncr2820710813>3.0. co;2-0.
7. Francis R. The application of binary logistic regression analysis on staff performance appraisal. Science J Applied Mathematics Statistics. 2017; 4:164-168. doi: 10.11648/j.sjams.20170504.15
8. Kleinbaum G.D, Klein M. Survival Analysis, A self-learning text. 3rd edi. Springer. 2012.
9. Kleinbaum G.D, Kupper L.L. Applied Regression Analysis and Other Multivariable Methods, 2nd edi. Boston: Duxbury. 1978;101 -102.
10. Kollias J, Elston CE, Ellis IO, Robertson JFR, Blamey RW. Early-onset breast cancer: histopathological and prognostic considerations. British J Cancer. 1997;75: 1318–1323. doi: 10.1038/bjc.1997.223
11. Kroman N, Jensen M. Factors influencing the effect of age on prognosis in breast cancer: population based study. British Med J. 2000. doi: 10.1136/bmj.320.7233.474
12. Meneses K, Azuero.A, Hassey L. Does Economic Burden Influence Quality of Life in Breast Cancer Survivors? Gynecol Oncol. 2012;124(3):437–443. doi:10.1016/j.ygyno.2011.11.038.
13. Nair S.K., Raj S. Cost of Treatment for Cancer: Experiences of Patients in Public Hospitals in India. Asian Pac. J Cancer Prev. 2014;5049-5054. doi: 10.7314/apjcp.2013.14.9.5049.

14. Rosen PP, Groshen S, Kinne DW, Norton L. Factors influencing prognosis in node-negative breast carcinoma: analysis of 767 T1N0M0/T2N0M0 patients with long-term follow up. J Clinical Oncol. 1993;11:2090–2100. doi: 10.1200/JCO.1993.11.11.2090.
15. Russo J, Frederick J. Predictors of recurrence and survival of patients with breast cancer. Am J Clinical Patho. 1987;88:123–131. doi: 10.1093/ajcp/88.2.123.
16. Seidman JD, Schnaper LA, Aisner SC. Relationship of the size of the invasive component of the primary breast carcinoma to axillary lymph node metastasis. Cancer. 1995;75:6571.doi:10.1002/10970142(19950101)75:1<65::aidcncr2820750112>3.0.co;2-b.
17. Smith JA, Gallager HS. Carcinoma of the breast: analysis of total lymph node involvement versus level of metastasis. Cancer. 1997;39:527–532. doi: https://doi.org/10.1002/1097-0142(197702)39:2<527::AIDCNCR2820390221>3.0.CO;2-N
18. Sun S.Y. Zhao Z. Risk Factors and Preventions of Breast Cancer. International J Biological Sci. 2017; 13(11): 1387-1397. doi: 10.7150/ijbs.21635
19. Veronesi U, Galimberti V, Zurrida S. Prognostic significance of number and level of axillary node metastases in breast cancer. The Breast. 1993;2:224–228. doi: https://doi.org/10.1016/0960-9776(93)90004-Y
20. Walter F.M, Rubin G. Symptoms and other factors associated with time to diagnosis and stage of lung cancer: A prospective cohort study. British J Cancer. 2015;S6–S13. doi: 10.1038/bjc.2015.30
21. Yavari K. Anti-Angiogenesis Therapy of Cancer Cells using^{153} SmBevasesomab. Emerging Sci. J. 2018;130-139. doi: http://dx.doi.org/10.28991/esj-2018-01136
22. Zare H. Effects of Salvia Officinalis Extract on the Breast Cancer Cell Line. SciMedicine J. 2019;25-29. doi: http://dx.doi.org/10.28991/SciMedJ-2019-0101-4

LIST OF ABBREVIATIONS

Abbreviations	**Meaning**
cm	Centimetre
ER	Estrogen Receptor
PR	Progesterone Receptor
APL	Above Poverty Line
BPL	Below Poverty Line
LN	Lymph Nodes
p	Probability of remission of cancer
β_i's	Regression Coefficients
x_i's	Explanatory Variables
logit(p)	Logarithm of the odds, where p is the probability
B	β_i's , Regression Coefficients
S.E	Standard Error
Wald	Wald Statistic
d.f.	Degrees of freedom
Sig.	Significant
CI	Confidence Interval
®	Reference Category
Viz.	Namely
HER2	Human Epidermal Growth Factor Receptor 2
AJCC	American Joint Committee on Cancer
T1	Tumour size with 1 centimetre