Available online at http://scik.org

J. Math. Comput. Sci. 2022, 12:144

https://doi.org/10.28919/jmcs/7322

ISSN: 1927-5307

MULTIVARIATE ANALYSIS OF THE BREAST CANCER PATIENTS IN NORTH-EAST INDIA: USING COX REGRESSION

SWAPAN BHATTACHARJEE*, SUROBHI DEKA

Department of Statistics, University of Cotton, Panbazar, Guwahati-781001, India

Abstract: Breast cancer is the most common cancer affecting women in India. Breast cancer can develop to any age groups of females and with the age the risk increases. The survival of breast cancer patients depends on early diagnosis, stages of the disease and treatment. Materials and Methods: In our study 462 breast cancer patients were included in study who were treated during 3 year period (January 2016 until December 2018) and were followed up to December 2019 at North East Cancer Hospital, Jorabat and State Cancer Hospital, Guwahati. Death reported was 75. The Kaplan-Meier was used for data analysis and in order to analyze the different covariates Cox Regression model is used. Results: Based on Kaplan- Meier mean survival days was 1132 and overall 3 year survival rate was 63%. The 3 year survival rate of stage III and stage IV was 60% and 20%. Conclusions: Our findings support the observation that those women with higher stages have less chance of survival. Moreover early detection of breast cancer may help to increase the survival rate of those women who are at risk.

Keywords: breast cancer; cox regression model.

2010 AMS Subject Classification: 62N01, 62N86.

*Corresponding author

E-mail address: swapanbhatta73@gmail.com

Received March 2, 2022

1. INTRODUCTION

Breast cancer is the major cause of mortality and morbidity among woman. Breast cancer has ranked number one cancer among Indian females with age adjusted rate as high as 25.8 per 100,000 women and mortality 12.7 per 100,000 women [1].Breast cancer accounts for 34% of all cancer cases among women in India[2]. According to Globocan 2018 report, in India the new cases of breast cancer in females was 162468 and deaths 87090[3] and Globocan 2020 report, in India the new cases of breast cancer was 178361 and deaths 90408 [4]. The five year survival rate of India ranged from 40 to 60% [5]. Breast cancer is the second most common cancer among women in Madras and survival rates at 1, 3 and 5 years were 80%, 58% and 48% [6]. The varieties of risk factors are associated with breast cancer are found to be age, diet, waist size, hip size, waist-hip ratio (WHR), body mass index, high-density lipoprotein cholesterol, triglyceride, more than three pregnancies, number of years of menstruation [7]. The treatment of breast cancer is Chemotherapy, Radiotherapy, Surgery (mastectomy and breast conserving surgery), hormone therapy and targeted therapy. M. Nadjib Bustan et al obtained that the hazard ratio of breast cancer patients of stage III and stage IV was 2.061[9]. The overall relative 5 year survival rate of breast cancer patients was found to be 62% and initial treatment(mastectomy, breast conserving surgery and neo-adjuvant therapy) was an independent predictor of poorer survival with hazard ratio 4.56[10]. Asrin Karimi et al showed that the survival time of breast cancer patients was 81 months and 5 year survival rate was 75% and patients with tumor metastasis had the survival rate of 37% [11]. Yuniar Farida et al showed in their paper that the probability of survival of breast cancer patients over two years is 73.7 % also showed that patients with tumor size less than 5 cm survive within two years is 3.497 times compared to breast cancer patients who have tumor size greater than 5 cm [12]. Mohamad Amin Pourhoseingholi et al showed that patients who were older than 45 years at diagnosis had an increased risk for death followed by greater tumor size and presence of pathologic distant metastasis [13].

The objective of this study was to determine the 3- year survival rate using Kaplan Meier approach and to study the covariates such as age at diagnosis, metastasis present or absent, histologic grades, tumor sizes using Cox Regression.

2. MATERIALS AND METHODS

It is a retrospective study and data were sourced from pathology report of the patients and hospital database record. In our study 462 breast cancer patients were included in study who were treated during 3 year period (January 2016 until December 2018) and were followed up to December 2019 at North East Cancer Hospital, Jorabat and State Cancer Hospital, Guwahati. Death reported was 75. Patients who were alive above 3 years were considered as censored observation. Clinical and pathological variables were age, clinical stage of the disease, metastasis of the tumor, grades and the site of metastasis.

Statistical Analysis:

Statistical analyses were performed using SPSS software; version 24 and in order to evaluate the proportional hazard assumption R programming software is used. Kaplan Meier curves were constructed for overall survival rate. The survival of a patient is referred to as the number of days from the date of diagnosis to the date of death. Cox multivariate analysis was also performed to analyze the different covariates using R programming. In order to check the proportional hazards assumption Scaled Schoenfled residuals test is used.

Results:

In this retrospective study, a total of 462 breast cancer patients were enrolled. Among all 382 (83.76%) were censored and 75(16.23%) patients were dead. The mean age of the patients at the time of diagnosis was 47.63±10.90 years. A total of 262 patients who were under age 50 years, out of which 224(85.5%) were censored and 200 patients who were above 50 years , out of which 163(81.5%) patients were censored and 37(18.5%) died (from Table 1). Based on Kaplan Meier method, mean survival days was 1132 and overall 3 year survival rate was 63%. The 3 year survival rate of stage III and stage IV was 60% and 20% (from Table 2). Prognostic factors

for breast cancer survival were analyzed by Cox multivariate analysis and the factors were clinical stage, metastases (localized or distant), tumor sizes and histological grades (from table 3). Proportional hazard assumption was met for all the factors. Based on risk of death for patients with stage III and stage IV is 7.3023 times higher than the patients with stage I and stage II (HR=7.3023;95% CI=3.636-14.67). Patients with metastases i.e distant metastases was 2.7439 times higher risk of death than patients with localized metastases (HR=2.7439; 95%CI=1.738-4.331). Patients with tumor size greater than 5cm 4.2955 times higher risk of death than patients with tumor size less than 2cm and 2-5 cm (HR=4.2955; 95%CI=2.551-7.233).

Table 1. The Number of Alive and Dead Breast Cancer Patients.

Factor		Total Observation	Censored (%)	Dead (%)
Age	<50	262	224(85.5)	38(14.5)
	>50	200	163(81.5)	37(18.5)
Clinical stage	Stage I	30	30(100)	0(100)
	Stage II	208	199(95.67)	9(4.3)
	Stage III	165	133(80.60)	32(19.39)
	Stage IV	59	25(42.37)	34(57.62)
Metastasis	Localized	354	316(89.26)	38(10.73)
	Distant	108	71(65.74)	37(34.25)
Site of metastasis	Bone	27	12(44.44)	15(55.55)
	Brain	37	25(67.56)	12(32.43)
	Lungs	37	22(59.45)	15(40.54)
	Liver	24	11(45.83)	13(54.16)
	Other	337	317(94.06)	20(5.93)
Tumor Size	<2cm	32	32(100)	0(100)
	2-5cm	250	231(92.40)	19(7.60)
	>5cm	180	124(68.88)	56(31.11)
Grades	WD	158	153(96.83)	5(3.16)
	MD	163	137(84.04)	26(15.95)
	PD	141	97(68.79)	44(31.20)

WD= Well Differentiated, MD= Moderately Differentiated, PD= Poorly Differentiated

Table 2. Survival rates of breast cancer patients according to possible prognostic factors.

Factor		Survival rate (%)		
		1 year	2 year	3 year
Age Groups	<50	95	80	60
	>50	90	80	63
Clinical Stage	Stage I			
	Stage II	99	93	91
	Stage III	93	80	60
	Stage IV	75	44	24
Metastasis	Localized	80	55	35
	Distant	98	98	84
Site of metastasis	Bone	80	49	9
	Brain	85	67	53
	Lungs	80	53	53
	Liver	78	56	
	Other	96	91	80
Tumor Size	<2cm			
	2-5cm	96	90	79
	>5cm	85	61	43
Grades	WD	98	95	83
	MD	94	77	59
	PD	85	67	50

Table 3. Multivariate Cox Regression Analyses of Breast Cancer Patients in North-East India

Factor	β	S.E	HR (95 % CI)	P value
Clinical Stage				
I,II			1	Reference
III,IV	1.9882	0.3558	7.3023(3.636-14.67)	0.0001
Metastases				
Localized			1	Reference
Distant	1.0094	0.2329	2.7439(1.738-4.331)	0.0001
Tumor Size				
<2cm, 2-5 cm			1	Reference
>5cm	1.4576	0.2659	4.2955(2.551-7.233)	0.0001

Checking for Proportional Hazard Assumption:

In order to check the Cox regression proportional hazard assumption using Schoenfeld residuals is used [8]. A small $p < 0.05$ value would indicate a violation of the proportionality assumption

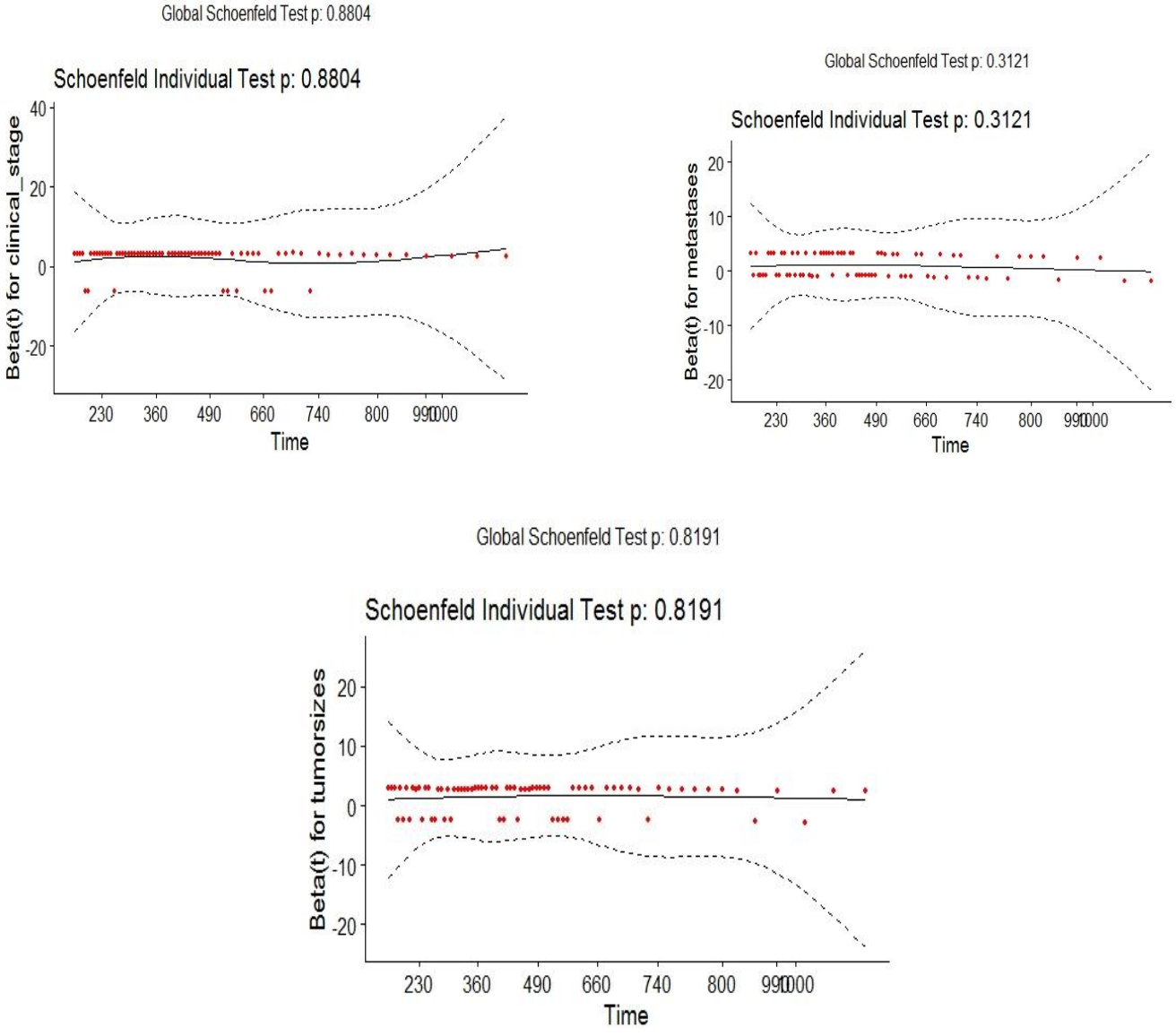

Fig 1: Breast cancer data, Schoenfeld residual plot for clinical stage, metastases and tumor sizes

3. RESULTS AND DISCUSSION

In this study our main aim was to study the different prognostic factors that affect the survival. Kaplan –Meier method was used to survival rate of different factors such as age groups, clinical stages (I, II, III, IV), metastases (localized or distant), and site of metastases, tumor sizes and different grades. Cox regression was used to investigate the prognostic factors. The proportionality of hazard assumption was checked by using Schoenfeld residuals. In our study it is found that patients with distant metastases, higher clinical stages, and greater tumor sizes have less survival rate.

CONFLICT OF INTERESTS

The author declares that there is no conflict of interests

REFERENCES

[1] S. Malvia, S.A. Bagadi, U.S. Dubey, S. Saxena, Epidemiology of breast cancer in Indian women, Asia Pac. J. Clin. Oncol. 13 (2017), 289-295.

[2] L. Subramanian et al., Breast Cancer Awareness in South India. Int. J. Sci. Stud. 6 (2018), 39-42.

[3] Globocan 2018: India Factsheet. International Agency for Research on Cancer. (2018).

[4] Globocan 2020: India Factsheet. International Agency for Research on Cancer. (2020).

[5] J. Sahwara, S. Bobdey, B.Ganesh, Breast cancer survival studies in India: A review. Int. J. Res. Med. Sci. 4 (2016), 3102-3108.

[6] C.K. Gajalakshmi, V. Shanta, R. Swaminathan, R. Sankaranarayanan, R.J. Black, A population-based survival study on female breast cancer in Madras, India. Br. J. Cancer. 75 (1997), 771-775.

[7] M.P. Antony, B. Surakutty, T.A. Vasu, M. Chisthi, Risk factors for breast cancer among India women: A case-control study, Nigerian J. Clin. Pract. 21(2018), 436-42.

[8] D. Schoenfeld, Partial residuals for the proportional hazards model, Biometrika 69 (1982), 551-555.

[9] M. Nadjib Bustan et al. Cox proportional hazard survival analysis to inpatients breast cancer cases, J. Phys.: Conf. Ser. 1028 (2018), 012230.

[10] M. Vahdaninia, A. Montazeri, Breast cancer in Iran: A survival analysis, Asian Pac. J. Cancer Prev. 5 (2004), 223-225.

[11] A. Karimi et al. Predictive factors of survival time of breast cancer in Kurdistan Province of Iran between 2006-2014: A cox regression approach, Asian J. Cancer Prev. 15(2014), 8483-8488.

[12] Y. Farida et al. Breast cancer survival analysis using cox proportional hazard regression and Kaplan Meier method. JTAM (Jurnal Teori dan Aplikasi Matematika). 5 (2021), 340-358.

[13] M.A. Pourhoseingholi, et al. Comparing Cox regression and parametric models for survival of patients with gastric carcinoma, Asian Pac. J. Cancer Prev. 8 (2007), 412-416.