MIKAILALSYS
Journal of Multidisciplinary Sciences

e-ISSN : 2987-2286
p-ISSN : 2987-3924

Index : Dimensions, Scilit, Crossref, Garuda, Semantic, Lens,Google Scholar,Base, etc

https://doi.org/10.58578/mikailalsys.v3i2.5365

Survival Analysis on Breast Cancer Incidences in Ikeja, Lagos Nigeria Using Kaplan Meier and Cox Proportional Hazard Techniques

Ajao Olutunde M.[1], Afariogun David A.[2], Ayeni Elizabeth O.[3], Egbe Kehinde O.[4]
Ajayi Crowther University Oyo, Oyo State, Nigeria
mo.ajao@acu.edu.ng; da.afariogun@acu.edu.ng

Article Info:

Submitted:	Revised:	Accepted:	Published:
Mar 15, 2025	Mar 30, 2025	Apr 11, 2025	Apr 16, 2025

Abstract

This study investigates breast cancer survival outcomes in relation to treatment types, using retrospective data sourced from the Oncology Department at Lagos State University Teaching Hospital (LASUTH). Employing survival analysis techniques—specifically the Kaplan-Meier method for estimating survival probabilities and the Cox Proportional Hazards model to examine the influence of demographic factors—the study measured survival time from diagnosis to either death or the last known follow-up. The event of interest was the patient's vital status. Results indicated that while breast cancer affects all genders, females are nearly twice as likely to be diagnosed as males. Age was also a significant determinant of survival, with older patients exhibiting a higher risk of mortality, as shown by increasing hazard ratios. These findings highlight the importance of early detection, particularly in younger individuals, as a critical strategy to enhance survival prospects.

Keywords: Breast cancer, Cox Proportional Hazard, Kaplan Meier, Survival analysis

INTRODUCTION

Cancer remains a global health challenge, transcending boundaries of ethnicity, gender, socio-economic background, and cultural context. Early detection and timely treatment are crucial, as delays can significantly reduce treatment success and patient survival outcomes. Broadly speaking, cancer encompasses a spectrum of diseases characterized by the unchecked proliferation and spread of atypical cells throughout the body (Diabate *et al.*, 2018). Among these, breast cancer represents one of the most prominent types, originating within the breast's anatomical structures—specifically, the lobules responsible for milk secretion and the ducts that convey milk to the nipple.

Breast malignancies are broadly classified into two categories: invasive and non-invasive. Invasive cancers have the potential to metastasize beyond the primary site, whereas non-invasive forms are localized. Historically, references to breast cancer were scarce until the Roman era, when the physician Celsus acknowledged surgical intervention as a viable option for early-stage cases. Contemporary literature on breast cancer in Africa frequently highlights the late-stage presentation of most patients, compounded by restricted access to awareness programs, screening tools, and adequate medical care (Vanderpuye *et al.*, 2017). While estimations regarding breast cancer incidence in Africa exist, they are often not grounded in real-time or population-specific data, leaving a knowledge gap in fully understanding the disease burden (Vanderpuye *et al.,* 2017).

In terms of predictive modeling, Breyer *et al.* (2018) explored various risk indicators for breast cancer in Southern Brazil, developing a multivariate logistic regression framework to estimate individual risk. Similarly, Tin *et al.* (2018) examined the disparities in breast cancer survival outcomes across ethnic groups in New Zealand, contributing to the growing body of evidence on demographic and systemic influences on survival.

Globally, cancer imposes substantial health and economic burdens across both affluent and resource-limited nations. According to international estimates, 14.1 million new cases and 8.2 million cancer-related deaths were recorded in 2012 alone (Torre et al., 2015). Breast cancer, in particular, is one of the leading causes of cancer mortality worldwide, ranking fifth overall and standing as the most frequent cause of cancer death among women across both industrialized and developing regions.

It is estimated that approximately 12.4% of women will face a breast cancer diagnosis during their lifetime. The likelihood of developing breast cancer rises with age, with most diagnoses occurring after the age of 50. The median age at diagnosis is 62 years. Moreover,

the disease exhibits substantial geographical variation in incidence, with higher rates documented in high-income nations compared to lower-income regions. This disparity is driven by a combination of demographic transitions, socio-economic development, improved diagnostic capabilities, increased disease awareness, and changes in reproductive and lifestyle factors.

In Nigeria, breast cancer poses a serious threat to the health and wellbeing of women, particularly those aged 50 years and older. Despite growing public health efforts, the incidence continues to rise annually. Greater awareness of breast cancer risk factors could empower individuals to take preventive actions and seek early diagnosis. For researchers exploring breast cancer survival trends using statistical methods such as the Cox proportional hazards model, relevant contributions can be found in the works of Ferraz and Moreira-Filho (2016), Tin *et al.* (2018), Breyer *et al.* (2018), and Pazvakawambwa and Embula (2017).

It is within this context that the present study is motivated—to support and extend ongoing research into breast cancer epidemiology, particularly within the Nigerian population, and to contribute to evidence-based strategies aimed at reducing the disease's impact.

MATERIALS AND METHODS

This study adopted a quantitative analytical approach, specifically utilizing a retrospective cohort design, given that the investigation relied on pre-existing secondary data. The target population encompassed all individuals diagnosed with breast cancer within Nigeria. For the purpose of empirical analysis, the sample comprised 2,182 documented cases of breast cancer from patients residing in Ikeja Local Government Area, Lagos State. These records were sourced from the Lagos State University Teaching Hospital (LASUTH) through the Department of Health Information Management, which maintains comprehensive clinical archives.

To analyze survival outcomes among breast cancer patients, the study employed both logit and probit binary logistic regression models in conjunction with the Kaplan-Meier estimator to evaluate survival probabilities over time. Furthermore, the Cox Proportional Hazards model was applied to assess the relationship between patient survival duration and selected demographic variables such as age, marital status, and occupation.

The estimates derived from the logistic regression analysis form the statistical foundation for understanding predictors of survival among breast cancer patients in the Nigerian context.

$$BCSA = e^{\alpha_0+\alpha_1 G+\alpha_2 A+\alpha_3 T+\alpha_4 MS} + \varepsilon_i \quad (1)$$

BCSA represents Breast Cancer

The parameters used in the study are defined as follows:

α_0= Constant

While$\alpha_1, \alpha_2, \alpha_3$ and α_4 are the coefficient of the independent variable G, A, T and Ms.

The independent variables include:

G represents gender

A represents Age

T represents time

MS represents Marital Status

The Kaplan-Meier model (Kaplan & Meier, 1958) estimator is given by

$$\hat{S}(t) = \prod_{i:t_i \leq t} \left(1 - \frac{d_i}{n_i}\right) \quad (2)$$

Where:

t_i denote time when at least one event happened.

d_i denote the number of events that happened at time t_i

n_i denote the number of individuals known to have survived up to time t_i (they have not yet had the death event or have been censored).

RESULTS AND DISCUSSION

The descriptive statistics analysis of the breast cancer incidence in Nigeria using Lagos State University Teaching Hospital (LASUTH) shows that the predominant gender

suffering from breast cancer in Nigeria are mostly female and the time of occurrence is mostly within 2 years of diagnosis.

The marital status analysis of breast cancer incidence reviews that most married women suffer from breast cancer disease and are likely to die. The treatment is mostly censored dead diagnosed asymptotically.

Table1: *Summary of Chi-square for Causal effect variables*

S/N	Variable interaction	χ^2_{cal}	Df	P value	P < 0.05	Remark
1	Censored * gender	0.364	2	0.833	0.833>0.05	Not significant
2	Censored*age	42.257	60	0.960	0.960>0.05	Not significant
3	Censored*time	30.416	11	0.001	0.001<0.05	Significant
4	Censored*marital status	0.311	3	0.958	0.958>0.05	Not significant
5	Censored*treatment	1.791	4	0.774	0.774>0.05	Not significant
6	Censored*diagnosis	6.661	2	0.036	0.036<0.05	Significant

Interpretation

i. The interaction term between censoring status and gender does not exhibit statistical significance, as evidenced by a p-value of 0.833, which exceeds the conventional threshold of 0.05.

ii. The effect of the interaction between censoring status and age is likewise not significant, with a p-value of 0.960, indicating no meaningful association.

iii. A statistically significant relationship is observed for the interaction between censoring status and time, given that its p-value of 0.001 falls well below the 0.05 cut-off.

iv. No significant interaction is found between censoring status and marital status, supported by a high p-value of 0.958.

v. The association between censoring status and treatment also lacks statistical significance, as shown by a p-value of 0.774.

vi. However, the interaction between censoring status and diagnosis reveals a statistically significant effect, with a p-value of 0.036, thus suggesting a potential influence on the outcome.

Logistic Regression Technique for Breast Cancer Survival Model

Table 2: *Logistic Regression*

Variable	B	Sig.	Exp(B)	df	Wald	Exp(B)>1	Remark
Gender	0.256	0.801	1.291	1	0.064	1.291>1	High risk
Age	0.004	0.760	1.004	1	0.093	1.004>1	High risk
Treatment	-0.187	0.010	0.829	1	6.632	0.829<1	Low risk
Marital status	0.289	0.651	1.336	1	0.204	1.336>1	High risk

From the table above shows the summarize logistics regression results of survival analysis. The model of breast cancer is shown as:

$$BC= e^{\alpha_0+\alpha_1 G+\alpha_2 A+\alpha_3 T+\alpha_4 MS} + \varepsilon_i \quad (3)$$

The model estimate of breast cancer is shown as:

$$BC= e^{1.903+0.256G+0.004A-0.187t+0.289MS} \quad (4)$$

The model of the estimate shows that there is a positive relationship among gender, age and marital status in the incidence of breast cancer. However, there is a negative relationship between treatment and breast cancer incidence in Nigeria.

Table 3: *Survival table at different event time*

event_at	removed	observed	censored	entrance	at_risk
0.0	0	0	0	726	726
1.0	456	437	19	0	726
2.0	118	111	7	0	270
3.0	59	53	6	0	152
4.0	38	32	6	0	93
5.0	19	17	2	0	55
6.0	10	10	0	0	36
7.0	15	13	2	0	26
8.0	7	7	0	0	11
9.0	1	1	0	0	4
10.0	1	0	1	0	3
11.0	1	1	0	0	2
12.0	1	1	0	0	1

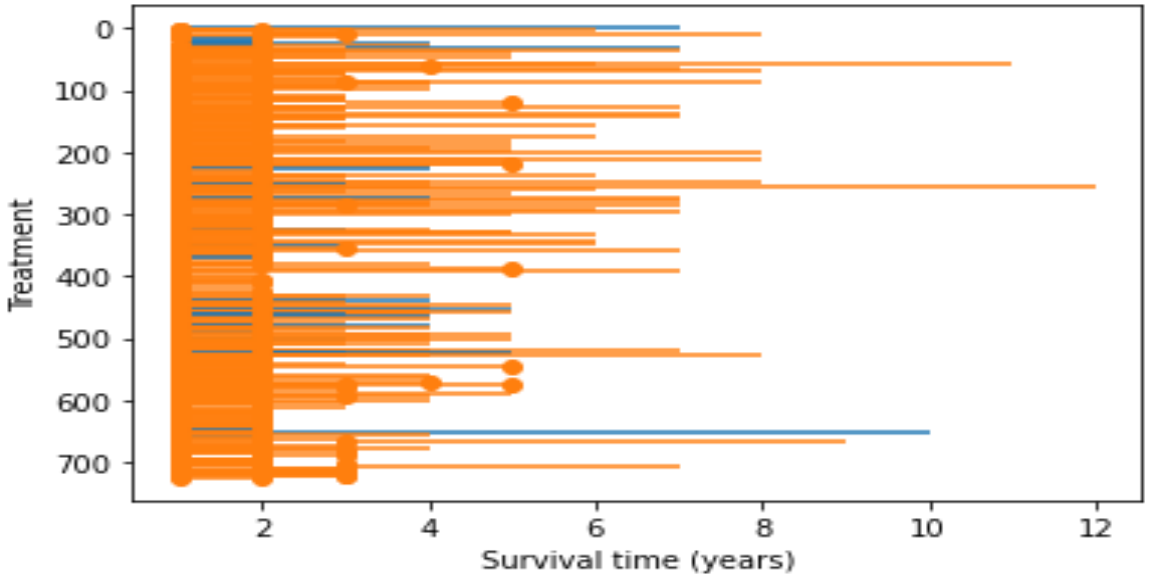

Fig.1: *Survival time based on treatment*

The survival analysis shows that there were more deaths within the first two years of diagnosis. Only very few survived longer than 4 years, with only one survivor after a 12 year follow up as shown in the table 3 and Fig 1 above.

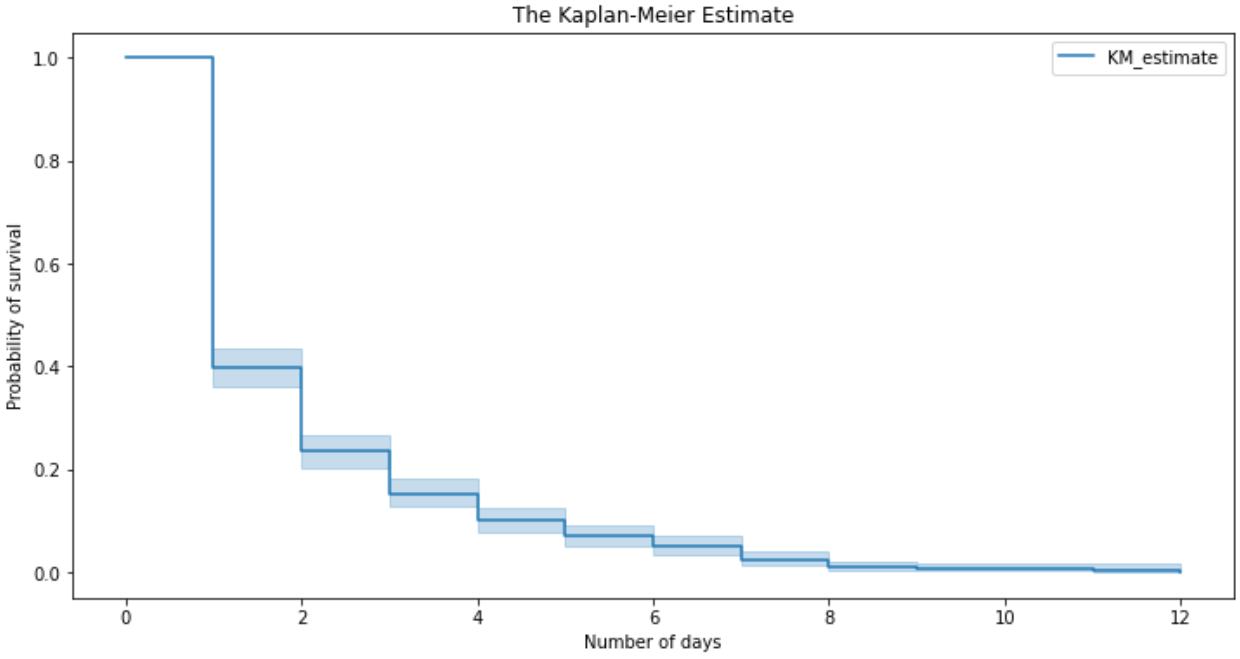

Fig 2: *Kaplan-Meier Estimate Of Survival Analysis*

The Kaplan-Meier estimate showed a proportional decline in probability of survival with respect to time as shown in the fig 2 above.

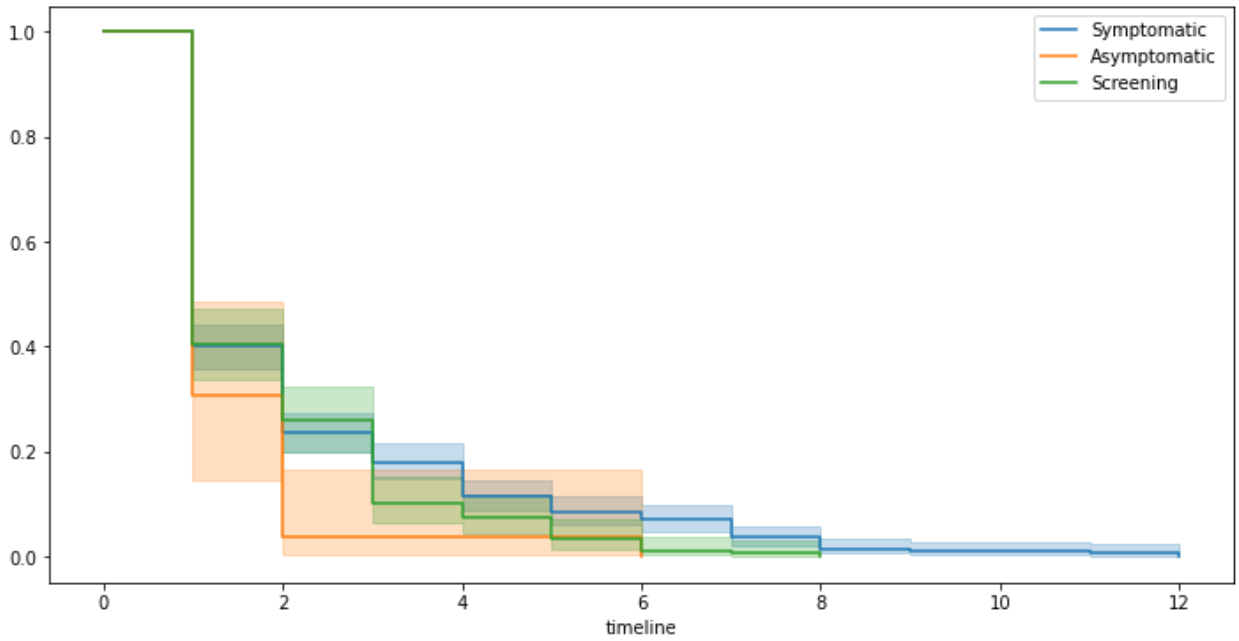

Fig 3: *Kaplan-Meier fit for survival time against treatment diagnosis groups*

Furthermore, the model fit for the different treatment and treatment diagnosis groups shows the different survival probabilities when different approaches were used. There is a higher probability of survival when patients were treated with radiotherapy as against when the cancer status had resulted to chemotherapy. Also, for the diagnosis, there seems to be a higher probability of survival when a symptomatic approach is followed as indicated in the fig 3 above.

Survival Analysis Model for Breast cancer

Table 4: *Cox Proportional Hazzard Model*

model	lifelines.CoxPHFitter
duration col	'survivaltime'
event col	'event'
baseline estimation	breslow
number of observations	726
number of events observed	683
partial log-likelihood	-3840.09
time fit was run	2022-06-19 10:41:05 UTC

	coef	exp(coef)	se(coef)	coef lower 95%	coef upper 95%	exp(coef) lower 95%	exp(coef) upper 95%	cmp to	z	p	-log2(p)
maritstatus	-0.09	0.92	0.18	-0.44	0.26	0.65	1.30	0.00	-0.49	0.62	0.69
age	-0.00	1.00	0.00	-0.01	0.00	0.99	1.00	0.00	-0.55	0.58	0.79
treatmt_Others	0.18	1.20	0.45	-0.71	1.07	0.49	2.92	0.00	0.40	0.69	0.54
treatmt_Palliative Care	-0.40	0.67	0.71	-1.80	1.00	0.17	2.72	0.00	-0.56	0.58	0.79
treatmt_Radiotherapy	-0.70	0.50	0.24	-1.16	-0.23	0.31	0.80	0.00	-2.92	<0.005	8.15
treatmt_Surgery	-0.34	0.71	0.09	-0.52	-0.16	0.60	0.85	0.00	-3.76	<0.005	12.53
gender_2	0.25	1.29	0.34	-0.41	0.92	0.66	2.51	0.00	0.75	0.46	1.13
gender_`	1.06	2.88	1.06	-1.03	3.14	0.36	23.18	0.00	1.00	0.32	1.65
tdiag_Screening	-0.06	0.94	0.22	-0.49	0.37	0.61	1.44	0.00	-0.28	0.78	0.35
tdiag_Symptomatic	-0.21	0.81	0.21	-0.62	0.20	0.54	1.22	0.00	-1.01	0.31	1.68

Concordance	0.59
Partial AIC	7700.18
log-likelihood ratio test	26.10 on 10 df
-log2(p) of ll-ratio test	8.12

The Cox-PH Model Results/Estimates shows the different coefficients for the different treatment types and different treatment diagnosis. The higher the hazard means more at risk of the event occurring.

i.e it shows that a one unit increase in surgery treatment means the baseline hazard will increase by a factor of exp(-0.34) = 0.66 % decrease.

Similarly, the values in the gender column are: [0 = male, 1 = female]. The value of the coefficient associated with gender, exp(0.25), is the value of ratio of hazards (Hazard Ratio) associated with being a diagnosed male patient compared to female diagnosis. This indicates the death risk is 1.29 times for patients who are male compared to female patients.

We use the fitted model in figure and table below to see how the survival changes as we change the covariate values. The plot partial effects on outcome method were used to see

how the survival varies for age group of 50, 60, 70 and 80-year-old patients compared to the baseline function. It clearly highlights that young patents have higher survival probabilities at any given instance of time compared to older patients.

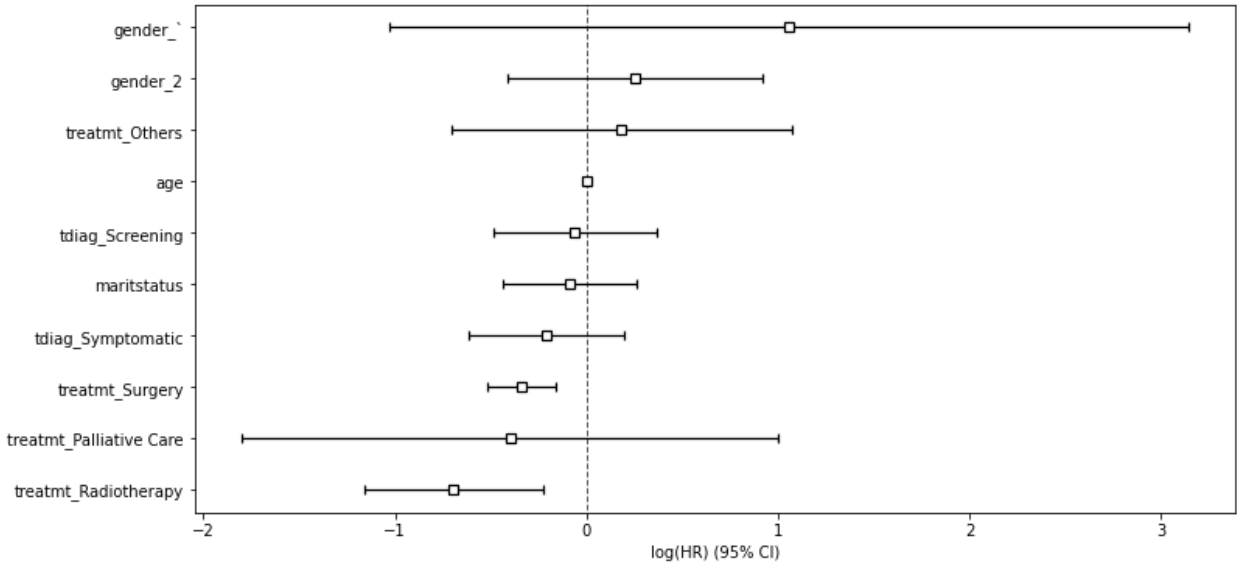

Fig. 5: *Cox Proportional Hazard Model Fit*

In terms of risk survival analysis of the disease based on the rate of occurrence and causal effect of breast cancer, the value of the Exp (B) measure with relative risk breast cancer disease based on the social economic characteristics and rate of occurrence of the disease. The results indicate that age, gender, marital status contributes relatively high risk to breast cancer disease in Nigeria.

Time and diagnosis of censored breast cancer is highly significant to breast cancer outbreak in Nigeria. Marital status, age, gender and treatment have similar pattern cancer censored in Nigeria based on the outcome of the study. The logistics model is significant and fit the equation of breast cancer estimation in Nigeria at the value of the cox-snell R square is $0.008 < 0.01$ at 1%.

The study statistical analysis of breast cancer incidence in Nigeria using Lagos State University Teaching Hospital (LASUTH) with the objective to evaluate the trend of breast cancer incidence, to examine the socio-economic characteristics of patients with breast cancer and to find out the relationship between cancer status and causal effects

CONCLUSION

The study reveals critical insights into demographic and treatment-related factors influencing the disease. The study confirms that breast cancer is significantly more prevalent among females, particularly within the age range of 48–49 years, highlighting age as a major risk factor. Marital status also contributes positively to the likelihood of developing breast cancer, with married women facing higher risks. While socioeconomic characteristics appear to have no statistically significant relationship with the incidence of the disease, the study found that timely and consistent treatment substantially reduces breast cancer occurrences, with an estimated 18.7% decline linked to treatment interventions. Also, the impact of treatment was found to be statistically significant, underscoring its importance in disease management. Overall, these findings emphasize the need for targeted awareness campaigns, early detection, and enhanced access to treatment for effective breast cancer control in Nigeria.

Recommendations

The following recommendation should be strictly adhered to, they are:

1. Seeking early diagnosis and treatment of the disease can go a long way in the reduction of Breast Cancer Incidence in Nigeria.
2. Prevention and health education and campaign should be embarked upon by government, individual, private and public sectors to avoid the occurrence scourges of Breast Cancer Incidence in the country.
3. Free health care services should be put in place by the government at Local, State and Federal level as a form of reducing the risk of Breast Cancer Incidence.
4. Donor agency and international organization should be made to actively participate in the control, prevention and management in Nigeria.

Acknowledgement: The authors appreciate the Department of Health Information Management of Oncology in Lagos State University Teaching Hospital (LASUTH) for providing data for the analysis in this research. We would also like to appreciate the reviewers and the editors for their suggestions which led to the success of this article.

Conflict of interests: The authors declare that there is no competing interest among them during the time of writing this paper.

Funding: Unfortunately, there is no funding for this research.

REFERENCES

Breyer, J. Z., Wendland, E. M., Kops, N. L., Caleffi, M., & Hammes, L. S. *Assessment of potential risk factors for breast cancer in a population in Southern Brazil. Springer Nature*, 169(1), Retrieved June 21, (2018), Pp 121-131.

Diabate, M., Coquille, L., & Samson, A. *Parameter estimation and treatment optimization in a stochastic model for immunotherapy of cancer.* (2018).

Ferraz, O. R., & Moreira-Filho, C. D. *Survival analysis of women with breast cancer: competing risk models*.Brazil:University of Campinas, (2016).

Kaplan, E. L. & Meier, P. *Nonparametric estimation from incomplete observations*. Journal of the American Statistical Association, 53(282), (1958), Pp 457–481.

Pazvakawambwa, L., &Embula, S. P. (2017). *Prevalence, trends and risk factors of Breast Cancer Mortality in Namibia*: 2000-2015. Windhoek: The University of Namibia. Retrieved September 28, (2018).

Tin, T. S., Elwood, M. J., Brown, C., Sarfati, D., Campbell, I., Scott, N., Lawrenson, R. *Ethnic disparities in breast cancer survival in New Zealand: which factors contribute?* BioMed Central. Retrieved September 28, (2018).

Torre, L.A., Bray, F., Siegel, R.L., Ferlay, J., Lortet-tieulent, J. & Jemal, A. *Global cancer statistics,* 2012. CA: A Cancer Journal for Clinicians, 65(2), (2015), Pp 87–108.

Vanderpuye, V., Grover, S., Hammad, N., Poojaprabhakar, Simonds, H., Olopade, F., & Stefan, D. C. *An update on the management of breast cancer in Africa*. BioMed Central, (2017), Pp 8-12.