

---

# Parametric Modeling of Survival Times Among Breast Cancer Patients in a Teaching Hospital, Osogbo

Phillip Oluwatobi Awodutire<sup>1, \*</sup>, Oladapo Adedayo Kolawole<sup>2</sup>, Oluwatosin Ruth Ilori<sup>3</sup>

<sup>1</sup>Department of Statistics, Federal Polytechnic of Oil and Gas, Bonny, Nigeria

<sup>2</sup>Department of Surgery, Ladoke Akintola University of Technology Teaching Hospital, Osogbo, Nigeria

<sup>3</sup>Department of Community Medicine, Ladoke Akintola University of Technology Teaching Hospital, Ogbomosho, Nigeria

## Email address:

phillip.awodutire@gmail.com (P. O. Awodutire), cullerwarley@yahoo.co.uk (O. A. Kolawole), oluwatosinnew@gmail.com (O. R. Ilori)

\*Corresponding author

## To cite this article:

Phillip Oluwatobi Awodutire, Oladapo Adedayo Kolawole, Oluwatosin Ruth Ilori. Parametric Modeling of Survival Times Among Breast Cancer Patients in a Teaching Hospital, Osogbo. *Journal of Cancer Treatment and Research*. Vol. 5, No. 5, 2017, pp. 81-85.

doi: 10.11648/j.jctr.20170505.12

**Received:** October 31, 2016; **Accepted:** November 21, 2016; **Published:** September 13, 2017

---

**Abstract:** In Nigeria, Breast Cancer is the most common malignancy among women. Unfortunately, many breast cancer patients present for treatment late. Using a parametric survival model to predict the survival times of patients and contribution of the prognostic factors, the study focused on the 1-year survival of breast cancer patients from the day of presentation. A total 89 women, who were diagnosed with breast cancer in which 32.56% reported early for treatment from 2009 to 2014, were recorded. Age, stage of presentation, average years of breastfeeding, neoadjuvant treatment offered, age at menarche and use of contraceptives were the variables used in the study. The predictive model that can be used to predict survival times of breast cancer patients was obtained. The results showed that stage at presentation is significant at 0.05 significance level.

**Keywords:** Survival Times, Parametric Model, Breast Cancer, Predictive, Prognostic Factor

---

## 1. Introduction

Cancer is known medically as a malignant neoplasm. It is a broad group of disease involving unregulated cell growth. In Cancer, cells divide and grow uncontrollably, forming malignant tumors and invading nearby parts of the body. The cancer may also spread to more distant parts of the body if not well taken care of at the early stage of formation. There are over two hundred different known cancers that affect humans but findings show that the commonest is Breast Cancer. [4]

Breast cancer occurs when cancer develops from breast tissue. Signs of breast cancer may include a lump in the breast, a change in breast shape, dimpling of the skin, fluid coming from the nipple, or a red scaly patch of skin. In those with distant spread of the disease, there may be bone pain, swollen lymph nodes, shortness of breath, or spinal cord paralysis.

In general, breast cancer is the most common malignancy among women in developed as well as some developing

countries, often being the second leading cause of cancer mortality after lung cancer [9].

Breast cancer affects women of all races without exception even though severity and survival rate are often diverse [10]. The breast cancer burden differs between countries and regions showing variations in incidence, mortality and survival rates. Early diagnosis of breast cancer is known to be vital not just in the treatment of the disease but also in determining prognosis [5]. Incidence rates are higher in the developed countries than in the developing countries. Incidence rates are also higher in urban areas than in the rural areas [4].

In Africa, breast cancer has overtaken cervical cancer as the commonest malignancy affecting women and the incidence rates appear to be rising. In the United Kingdom, where the age standardized incidence and mortality is the highest in the world, the incidence among women aged 50 approaches two per 1000 women per year, and the disease is the single commonest cause of death among women aged 40-50, accounting for about a fifth of all deaths in this age group.

There are more than 14 000 deaths each year, and the incidence is increasing particularly among women aged 50-64, probably because of breast screening in this age group [6].

In Nigeria for example, incidence rate has increased from 13.8-15.3 per 100,000 in the 1980s, to 33.6 per 100,000 in 1992 and 116 per 100,000 in 2001. These increases in incidence are due to changes in the demography, socio-economic parameters, epidemiologic risk factors, late reporting and unawareness of the disease. While mortality rates are declining in the developed world (Americas, Australia and Western Europe) as a result of early diagnosis, screening, and improved cancer treatment programs, the converse is true in the developing countries [2].

Factors that increase the incidence of breast cancer include early menarche and late menopause, null parity and an increased age at the first birth. Also, family history of breast cancer before the age of 40, bilateral breast cancer, ovarian cancer, and male breast cancer are also associated with breast cancer [7]. Many of these factors cannot be prevented. Therefore, early screening is still the best way to control the disease and configures one of the strategies to increase the survival rate of the women. Survival refers to the occurrence of a specific event of interest, starting from an initial time until a final time, for example, from the diagnosis of breast cancer to the death. Thus, the analysis of survival means the likelihood of survival from zero time until the final time considered. Knowing the factors that contribute to increased survival rates among women with breast cancer may help define early detection actions, and improve treatment and care proposals, in the nursing area as well as in all the areas of health.

## 2. Breast Cancer in Nigeria

In the North-Western geopolitical zone of Nigeria, cancer of the breast was second to cancer of the cervix, while at University College Hospital (UCH), Ibadan (situated in the South-Western geopolitical zone of Nigeria) it was the leading malignancy among women. In the North-Central geopolitical zone, breast cancer constituted 22.41% of new cancer cases registered in 5 years and accounted for 35.41% of all cancers in women.

Unfortunately there is paucity of data and sparse literature review on the trends of breast cancer in Nigeria due to few existing cancer registries most of which are either hospital-based or pathology based instead of the preferred population-based cancer registries. Also, in low resource countries, hospital-based cancer registry has been serving as a fundamental source of information on cancer [3].

In Nigeria, breast cancer was characterized by late clinical presentation and in advanced stage of the disease, when only chemotherapy and palliative care could be offered, and therefore associated with high mortality [1].

In Nigeria, as indeed in many developing countries, a combination of poor health education, poverty and a high patronage of non-orthodox healing practices among the populace contribute to late presentation of breast cancer in many hospitals with attendant high number of metastatic

disease and poor disease survival. This is worsened by the commonly encountered non-adherence to treatment schedule among the patients. The burden of caring for these large numbers of patients in a low resource country is enormous [2]. The purpose of this study is to derive the model that can predict survival times of breast cancer patients and determine which prognostic factor.

## 3. Related Works

According to Aalen [11], the parametric model is underused in medical research and deserving of more attention. Elvan [16] compared five survival models using breast cancer registry data from Ege University Cancer Research Center. A survival analysis was conducted by using Weibull, Gamma, Gompertz, Loglogistic and Lognormal distribution of parametric models. In the analysis of the survival periods using parametric models, the age variable is taken as the covariate. To determine the best model among parametric models, Akaike Information Criteria (AIC) was exploited. The results of the study revealed that the survival model found by the Gompertz distribution was the most appropriate one.

Zare [8] worked on modelling of breast cancer prognostic factors using a parametric log-logistic model in Fars province, Southern Iran. The log-logistic model was employed as the best parametric model which could explain survival times. The hazard rates of the poor and the medium prognosis groups were respectively 13 and 3 times greater than in the good prognosis group. Also, the difference between the overall survival rates of the poor and the medium prognosis groups was highly significant in comparison to the good prognosis group.

Vallinayagam [15] compared the performance of the common parametric models to a breast cancer survival data and found out that lognormal model is better than other models after comparison using the likelihood ratio test.

Baghestani *et al.* [12] used the weibull parametric model to analyse the survival time of breast cancer patients that were treated at the cancer research centre in Shadid Beheshti University of Medical Sciences and found out that the one year overall survival rate was 0.98.

## 4. Materials and Method

### 4.1. Study Site

Ladoke Akintola University of Technology Teaching Hospital, Osogbo, the capital city of Osun state was established in 1991. It is located in Osun state in the South-Western part of Nigeria. It is jointly owned and funded by two states (Oyo state and Osun state)

### 4.2. Study Design

The study was a descriptive study conducted on breast cancer cases from 2009 to 2014. The study was carried out at the general surgery ward of the hospital. Four parametric survival

models were used to model survival times of the breast cancer patient (a regression model) and determine the effect of the prognostic factors towards the survival times of the patient.

**4.3. Data Collection**

Clinical data from eighty-nine selected breast cancer patients were retrieved from their case-note files at the Ladoke Akintola University of Technology Teaching Hospital, Osogbo. The survival times of the patients were retrieved from their files, right censored at one year of diagnosis. The survival time recorded is the time from the day of report (admission) of the patient till the day of last contact (death, alive or loss to follow up). Also, variables (factors) that might influence the survival time of the patient were also recorded. The variables retrieved for this research are Age of Patients at report (Age), Age of Patient at menarche (menarche), Use of Contraceptives (contraceptives), Average years used for breastfeeding (breastfeeding), Stage of tumor development at the point of report (detection) and the use of Neoadjuvant Therapy (neoadjuvant).

**4.4. Methods**

Parametric methods assume that the underlying distribution of the survival times follows certain known probability distribution. Popular ones include the exponential, weibull, log-logistics and lognormal distributions. The description of the distribution of the survival times and the change in their distribution are usually estimated using maximum likelihood method. Survival estimates obtained from parametric survival models typically yield plots more consistent with the theoretical survival curve. Also, efficiency and completeness of estimates are the main appeals of using the parametric approach.

*Accelerated Failure Time Model (AFT)*

The AFT is a predictive (regression) model in which the survival times (time to event) depends on the covariates. Let  $T_i$  be a random variable denoting the survival times of the  $i^{th}$  individual in the sample,

$x_{ip}, p= 1, \dots, k$  be the values of  $k$ -covariates for the same individual and

$i = 1, \dots, n$  where  $n$  is the sample size, the model is

$$\ln T = \delta + x' \beta + \varepsilon \tag{1}$$

where  $\varepsilon$  is said to follow a particular distribution,  $\beta$  are the estimates of the covariates  $x$  and  $\delta$  is the intercept of the model.

Estimates of the parameters of the AFT model can be derived by using the Maximum Likelihood Estimation (M. L. E) method, involving optimization technique. The maximum likelihood function is of the form

$$L = \prod_{i=1}^r f(t_i/x, \beta) + \prod_{r+1}^n s(t_i^+/x, \beta) \tag{2}$$

which equivalently is

$$L = \prod_{i=1}^r g(x/\beta) f_o(Z(t_i)) + \prod_{r+1}^n s_o(Z(t_i)) \tag{3}$$

where  $g(x/\beta) = \exp(-x\beta)$  and  $Z(t) = t \exp(-x\beta)$ .

Also  $f_o(Z(t))$  and  $s_o(Z(t))$  are the baseline probability density function and baseline survival function respectively.

**4.5. Data Analysis**

The retrieved data was analyzed using four common existing survival models (exponential, weibull, lognormal, loglogistic) with the aid of R-statistical software. In this study, Akaike Information Criterion (AIC) was used to measure the goodness of statistical models fitness (determine the considerably best model). The smaller the AIC, the better it is [17]. AIC for the model used in this study has been calculated according to the following equation

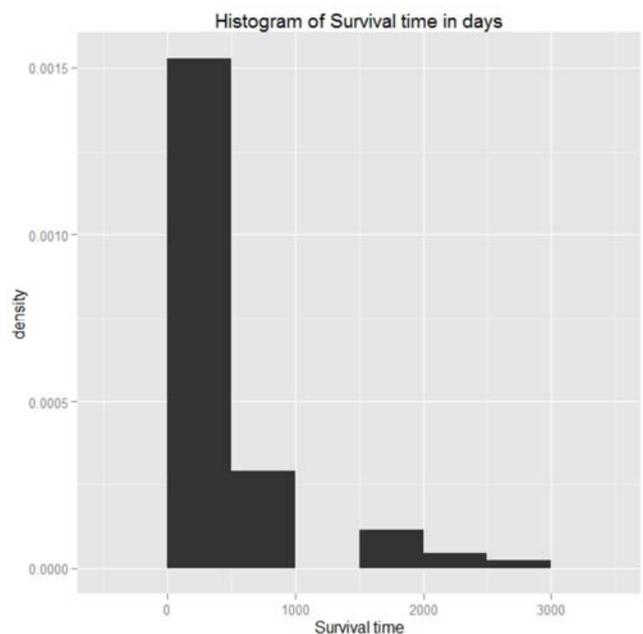
$$AIC = -2 \log(\text{likelihood}) + 2k \tag{4}$$

in which  $k$  is the number of parameters in the model.

Also, from the model with the lowest AIC, the prognostic factor was considered significant if the p-value is less than 0.05 (i.e  $p < 0.05$ ) and if otherwise, the prognostic factor is insignificant according to the result of the study.

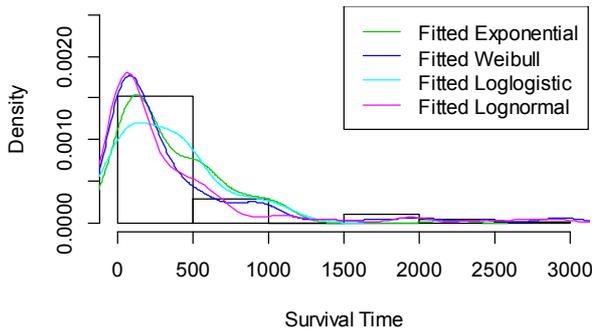
**5. Results**

A total 89 women were diagnosed with breast cancer in which 32.6% reported early for treatment. The mean age at presentation is 50.29 and 46.07% used contraceptives. Figure 1 is a chart showing the survival time distribution of the patients. The distribution is skewed to the left as could be seen from the histogram. The skewness of this survival time distribution is pointing to the fact that as the survival time increases, the number of survivors decreases. Figure 2 displays the histogram showing the distribution curves of the fitted models under consideration which shows that the data is well fitted by the distributions.



**Figure 1.** Histogram of the survival times of Breast cancer patients.

**Histogram of Survival Time and the Fitted Distributio**



**Figure 2.** Histogram of survival time and the fitted distribution.

Using the four survival models to model the survival times of the patients, the tables 1-4 show the results of the estimates of parameter and p-value of each parameter. Comparing the AIC of the derived models, table 5 shows that lognormal clearly demonstrates superiority over these other models due to its lowest AIC. Using the lognormal model the result of the study showed that the stage of reporting (detection) was significant.

**Table 1.** Analysis of the survival times of breast cancer patients using loglogistic survival model.

Parameters	Value	Std. Error	z	p
(Intercept)	-1.5145	3.0310	-0.500	6.17e-01
Age	0.0382	0.0319	1.199	2.31e-01
Menarche	-0.0886	0.1341	-0.661	5.09e-01
Breastfeed	0.7485	0.6288	1.190	2.34e-01
Contraceptive	1.1620	0.6551	1.774	7.61e-02
Detection	1.5490	0.7115	2.177	2.95e-02
Neoadjuvant	1.1762	0.6844	1.719	8.57e-02
Log(scale)	0.4672	0.1188	3.932	8.41e-05

**Table 2.** Analysis of the survival times of breast cancer patients using lognormal survival model.

Parameters	Value	Std. Error	z	p
(Intercept)	-1.2598	3.1709	-0.397	6.91e-01
Age	0.0406	0.0325	1.250	2.11e-01
Menarche	-0.1018	0.1394	-0.731	4.65e-01
Breastfeed	0.6021	0.6237	0.965	3.34e-01
Contraceptive	1.1087	0.6550	1.693	9.05e-02
Detection	1.4720	0.7163	2.055	3.99e-02
Neoadjuvant	1.3543	0.6995	1.936	5.28e-02
Log(scale)	1.0062	0.1076	9.356	8.30e-21

**Table 3.** Analysis of the survival times of breast cancer patients using exponential survival model.

parameter	Value	Std. Error	z	p
(Intercept)	1.8867	1.7423	1.083	0.278865
Age	0.0347	0.0157	2.208	0.027258
Menarche	-0.0562	0.0690	-0.815	0.415135
Breastfeed	0.1807	0.3364	0.537	0.591173
Contraceptive	0.6521	0.3009	2.168	0.030184
Detection	1.1660	0.3321	3.511	0.000446
Neoadjuvant	0.6491	0.3195	2.032	0.042166

**Table 4.** Analysis of the survival times of breast cancer patients using weibull survival model.

Parameter	Value	Std. Error	z	p
(Intercept)	-0.4937	3.2719	-0.151	8.80e-01
Age	0.0431	0.0318	1.355	1.76e-01
Menarche	-0.1135	0.1379	-0.823	4.10e-01
Breastfeed	0.6967	0.6458	1.079	2.81e-01
Contraceptive	1.1654	0.6488	1.796	7.25e-02
Detection	1.6083	0.6962	2.310	2.09e-02
neoadjuvant	1.2938	0.6695	1.932	5.33e-02
Log(scale)	0.7781	0.1171	6.643	3.07e-11

**Table 5.** Table showing comparison of AIC of the survival models.

Model	Loglikelihood	AIC
Exponential	-374.3	762.6
Lognormal	-337.9	691.8
Loglogistic	-338.7	693.4
Weibull	-341.9	699.8

## 6. Discussion

Breast cancer is a clinically heterogeneous disease. This study showed that lognormal model is considerably the best fitted life time model as supposed to weibull distribution gotten in a study done in Pakistan. Survival rate are worse when compared to those in older women and multivariate analysis has shown younger age to be independent predictor of adverse outcome in a study done by Carey *et al* [14]. However, age seems not to be significant predictor of survival in this new finding. This may be due to the fact that cancer can develop both in the young and old and therefore there should be no age limit to when self-breast examination should be instituted. Self-breast examination should be encouraged even right from adolescence. However, in this study, early detection of breast cancer is a good prognostic factor for survival; this was in line of what was discovered in a study done in Spain, which collaborated the fact that stage of presentation (detection) was a good predictor of survival times for breast cancer patients [9].

## 7. Conclusion

In this study, four parametric survival models were used to model the survival times of breast cancer patient in a teaching hospital in Osogbo. The model is a predictive model of survival times, considering the prognostic factors. The AIC was used to compare the models. Out of the four models used in this study, lognormal survival model which has the lowest AIC value was the considerably best model. In the lognormal model, stage of presentation (detection) of the breast cancer is found to be statistically significant with the survival experience while other factors are insignificant.

## Recommendations

Self-breast examination should be advocated. If breast cancer is detected early, survival time is improved. This is because once a lesion is detected early, preventive or curative

measures can be instituted before further spread of the tumor. However, some of the known modifiable risk factors for breast cancer can be engineered such as prevention of obesity, avoidance of consumption of fatty foods etc. Also, we advocate for the government to promote breast cancer awareness practices (especially the need for early reporting of breast cancer) among females especially adult females. The government should further help in making cancer registries functional in Nigeria to enhance proper documentation of cancer patients.

---

## References

- [1] Adetifa, F. A. and Ojikutu, R. K. (2009): Prevalence and Trends in Breast Cancer in Lagos State, Nigeria. *African Research Review*, 3 (5).
- [2] Adisa, A. O., Arowolo, O. A., Akinkuolie, A. A., Titiloye, N. A., Alatise, O. I., Lawal, O. O. and Adesunkanmi, A. R. K. (2011): Metastatic breast cancer in a Nigerian tertiary hospital. *African Health Sciences*, 11 (2).
- [3] Afolayan, A. (2012): Breast cancer trends in a Nigerian population: an analysis of cancer registry data. *International Journal of Life Science and Pharma Research*, 2 (3).
- [4] McPherson, K., Steel, C. M., Dixon, J. M. (2000): ABC of Breast Diseases, *BMJ* vol 21.
- [5] El Saghir N. S., Adebamowo C. A., Anderson B. O., Carlson R. W., Bird P. A., Corbex M., Badwe R. A., Bushnaq M. A., Eniu A., Gralow J. R., Harness J. K., Masetti R., Perry F., Samiei M., Thomas D. B., Wiafe-Addai B., Cazap E. (2011): Breast cancer management in low resource countries (LRCs): consensus statement from the Breast Health Global Initiative. *Vol 20 Suppl 2: S3-S11*.
- [6] Parkin, D. M., Bray, F., Ferlay, J., Pisani, P. (2002): Global Cancer Statistics, CA; a Cancer Journal for Clinicians 2005 Mar; 55 (2): 74-108.
- [7] Luciana Martins da Rosa<sup>1</sup>, Vera Radünz (2012): Survival Rates to Woman with Breast Cancer: Review Text Context Nursing, Florianópolis, 2012 Oct-Dec; 21 (4): 980-9.
- [8] Najaf, Z., Marzieh D., Abass R. (2012): Modeling of Breast Cancer Prognostic Factors using a Parametric Log-Logistic Model in Fars province, Southern Iran. *Asian Pacific Journal of Cancer Prevention*, Vol 13, 2012 1537.
- [9] Montesarrat Rue, Sandra Lee, Ester Vilaprinyo, Montserrat Martinez-Alonso, Misericordia Carles, Rafael Marcos-Gragera, Roger Pla and Josep- Alfons Espinas: Effectiveness of Early Detection of Breast Cancer Mortality Reduction in Catalonia (Spain). *BMC Cancer* 2009; 9: 326, doi: 10.1186/1471-2407-9-326.
- [10] Ojewusi Ayoola A., Obembe Taiwo, Aruogun Oyedunni S and Olugbayela Tunde (2016): Breast Cancer Awareness, Attitude and Screening Practices in Nigeria; a systematic review. 7 (2), pp. 11-25.
- [11] Aalen, O. O. (2000): Medical statistics-no time for complacency. *Statistical Methods In Medical Research* 9, 31-40.
- [12] Baghestani A. R., Moghaddam S. S., Majd H. A, Akbari M. E, Nafissi N., Gohari K. (2015): Survival Analysis of Patients with Breast Cancer using Weibull Parametric Model. *Asian Pacific Journal of Cancer Prevention*, article 85, vol 16, issues 18, pages 8567-8571.
- [13] Usman M., Dikko H. G., Bala S., Gulmbe S. U (2014): An Application of Kaplan-Meier Survival Analysis using Breast Cancer DATA. *Sub-Saharan African Journal of Medicine/ vol1/Issue3*.
- [14] Carey K. Anders, Rebecca Johnson, Jennifer Litton, Marianne Phillips and Archie Bleyer (2009): Breast cancer before age 40 years. *Semin Oncol.* June; 36(3): 237-249. doi: 10.1053/j.semincol.2009.03.001
- [15] Vallinayagam V., Prathap S. and Venkatesan P. (2014): Parametric Regression Models in the Analysis of Breast Cancer Survival Data; *International Journal of Science and Technology*. Vol 3, No. 3.
- [16] Elvan A. H (2010): Comparison of Five Survival Models; Breast Cancer Registry Data from Egw University Cancer Research Center, *International Biometric Society Journal*.
- [17] Collett, D. (2003): *Modelling Survival Data in Medical Research*, Second Edition; Chapman & Hall/ CRC Press, Boca Raton, FL.