International Journal of Innovation in Science and Mathematics
Volume 2, Issue 6, ISSN (Online): 2347–9051

Comparative Analysis of Parametric Models on Survival of Breast Cancer Patients in Ghana

Alice C. Mensah
Mathematics and Statistics Department,
Accra Polytechnic, Accra
Email: alicecabakah@yahoo.com

Dorothy Anima Frimpong
Computer Science Department,
Accra Polytechnic, Accra

Ebenezer Okyere
Bank of Ghana, Research Department,
Accra

Abstract – **Survival analysis methods that measure the risk of death or progression of a disease provide predictions that help clinicians to estimate trends in their patient outcomes. The objective of the study was to determine the survival pattern of breast cancer patients, using the parametric modeling strategies. Five parametric models-exponential, Weibull, Lognormal Gamma and Llogistic—were applied to the real life data which consisted of 1022 women diagnosed with breast cancer between 1st January 2002 and 31st December 2008. Survival time was calculated from the date of the diagnosis of breast cancer to the date of death or, if alive, at 31 December 2011. Using the log likelihood method and the Akaike information criterion (AIC) the gamma model was found to be the best-fitted model for predicting survival following a diagnosis of breast cancer. Several covariates-including size of tumour, tumour grade; stage at diagnosis; axillary node involvement; Body Mass Index (BMI) and Age (age of the patient in years)—were included in the parametric model to predict factors associated with future mortality. Size of tumour, stage at diagnosis and Body Mass Index (BMI) were found to be significant variables associated with mortality of breast cancer patients.**

Keywords – **Survival Analysis, Breast Cancer, Parametric Model, Akaike Information Criterion.**

I. Introduction

Survival analysis pertains to a statistical approach designed to take into account the amount of time an experimental unit contributes to a study period, or the study of the time between entry into observation and a subsequent event, (Allison, 1995). The analysis of survival data involves a collection of statistical procedures for which the outcome variable of interest is time until an event occurs. Data that has censored observations are said to be incomplete and the analysis of such data requires special techniques. There are basically three methods that could be used in the analysis of survival data namely: fully parametric, non-parametric and semi-parametric.

Fully parametric methods assume the knowledge of the distributions of the survival time e.g. exponential, weibull, loglogistic, log-normal, gamma and gompertz. Non-parametric models make no assumptions of the distribution of the survival times, e.g. the Kaplan-Meier estimators. The semi-parametric models assume a parametric form for the effects of the explanatory variables but make no assumptions on the distributions of the survival times.

Breast cancer in women is a major health burden. It is the most common cancer diagnosed in women and also the most common cause of cancer death among women in both high recourse and low recourse countries, (Bray *et al,* 2004). The incidence, mortality and survival rate in different parts of the world vary from 4 to 10 fold. It is the 2nd leading cause of death in women worldwide,(Okobia *et al*,2006). The burden of the disease both in developed and developing countries is increasing and if no action is taken it will go beyond our control. According to IARC 1.5 million new cases of Breast cancer was diagnosed in 2002, and among them approximately 411,000 died. Based on current estimate of an average annual increase in incidence ranging from 0.5% to 3% per year, the projected incidence increase in 2010 will be 1.4-1.5 million, (Smith,*et al,* 2006). In Ghana data on breast cancer is scant. However, the disease is a common cause of hospital admissions and mortality among Ghanaian women, (Wiredu and Armah, 2006). Reported clinical studies from Ghana and other communities in sub Saharan Africa indicate that breast cancer in indigenous black African population is often severe with unfavourable prognostic features, (Gukas et al, 2005; Amir et al, 1994). Some of these features include young age at presentation, advanced stage at diagnosis, large tumour size, high grade histologic subtypes and low rate of hormone receptor positivity, (Gakwaya et al,2008; Clegg-Lamptey and Hodasi, 2007; Mbonde et al, 2000; Mbonde et al, 1998; Amir et al, 1997 and Hassan et al, 1992. This study aimed to com pa re the results of the survival analysis of the breast cancer patients by using exponential, weibull, log-normal, gamma and log-logistic parametric models. Before conducting the parametric modelling, the patients were divided in to two groups as 49-and-lower and above 49years.

II. Methods

This is a retrospective cohort study comprised all Ghanaian women who reported at the Breast Clinic of the Korle_bu Teaching Hospital (KBTH) or its cancer centre between January 2002 and December 2008. The 1022 cases were sampled from the Records Section of the Cancer centre of the KBTH. Cases were required to have histologically proven breast cancer consisting of invasive ductal carcinoma, invasive lobular carcinoma, ductal carcinoma in-situ and lobular carcinoma in-situ. Survival time was calculated from the date of the diagnosis of breast cancer to the date of death or, if alive, at 31st December 2011. Data on patients were extracted from the abstract of hospital records. The patients were divided into two with respect to their age as 49 and lower and above 49years. The log rank test was applied to compare the survival curves of the two age groups obtained by Kaplan Meier method. The survival analysis was conducted by using exponential, weibull, log-normal, gamma and log-logistic parametric models. Covariates – including size of tumour, tumour grade; stage; axillary node; BMI and Age (age of the patient in years)- were included in the parametric model..

All models were fitted to the dataset. The maximum likelihood method and the Akaike information criterion (AIC), were used to assess the best parametric model fitted to the breast cancer dataset, which gives a measure of the goodness of fit of an estimated statistical model.
It is given by:

AIC= - 2log (likelihood) + 2(p + k), where p is the number of parameter, and k is the number of parameters in the distribution. Statistical models with lower AIC are preferred.

Table 1: The hazard functions, probability density functions and survival functions for the parametric distributions (Ata *et al.*, 2008)

Distribution	Parameter	h(x)	f(x)	s(x)
Weibull	α >0 λ> 0	$\lambda\alpha x^{\alpha-1}$	$\lambda\alpha x^{\alpha-1}\exp(-\lambda x^{\alpha})$	$\exp(-\lambda x^{\alpha})$
Log - logistic	α ,λ >0	$\frac{\alpha x^{\alpha-1}\lambda}{1+\lambda x^{\alpha}}$	$\frac{\alpha x^{\alpha-1}\lambda}{[1+\lambda x^{\alpha}]^2}$	$\frac{1}{1+\lambda x^{\alpha}}$
Log normal	α >0	$\frac{f(x)}{s(x)}$	$\frac{\exp\left[-\frac{1}{2}\left(\frac{\ln x-\mu}{\sigma}\right)^2\right]}{x(2\pi)^{1/2}\sigma}$	$1-\phi\left[\frac{\ln x-\mu}{\sigma}\right]$
Gamma	α,β, λ >0	$\frac{f(x)}{s(x)}$	$\frac{\alpha\lambda^{\beta}x^{\alpha\beta-1}\exp(-\lambda x^{\alpha})}{\Gamma(\beta)}$	$1-\mathrm{I}[\lambda x^{\alpha}\beta]$
Exponential	Λ>0	$\frac{f(x)}{s(x)}$	λexp(- λt)	exp(- λt)

This research has been assessed and approved by the School of Allied Health Sciences, University of Ghana Ethics and Protocol Review Committee with identification number: SAHS – ET./AA/1A/2013-2014. SAS version 9.0 (Cary, NC 27513 USA) was used for the analysis. P - Value < 0.05 was considered significant.

III. Results

The age for the breast cancer patients ranged from 20 – 92 years with a (Mean ± SD) value of 47.97 ± 11.77 years and median of 47 years. The mean value for Body Mass Index (BMI), Tumour Size (TS) and Grade (GR) were, 28.27, 5.57 and 2.03 respectively. BMI values ranged from 15.24 – 76.1, that of TS was 0.5 – 35cm and GR values ranged from 1 – 3. The site of the tumour on presentation was the left breast in 46.57% of cases and the right breast in 53.43% of cases. Lump size on presentation (revealed by ultrasound, mammography and/or clinical palpation) in 71.28% of the women was 2-5 cm on first diagnosis, 4.46% was less than 2 cm and in 24.26% of the women was greater than 5 cm on diagnosis. Axillary lymph node involvement was found in 90% of the women diagnosed with breast cancer when they were first seen by a physician; 89.20% had axillary lymph nodes of more than 25% involvement. Data relating to the clinical stages of breast cancer on first diagnosis showed the highest stage of presentation was stage III, accounting for 47.16% of the cases. In all, 52.94% of the diagnosis was presented in the advanced stage (stage III and IV). There were 390 cause-specific deaths as a result of breast cancer, 55.97% of the deaths occurred in the age group of 49years and below. Application of the Kaplan – Meier method to the data resulted in the following information: The mean survival time was 4.59 years (55.13months). The 5 - year overall survival was 47.9%. The survival curve is shown in Fig.1.

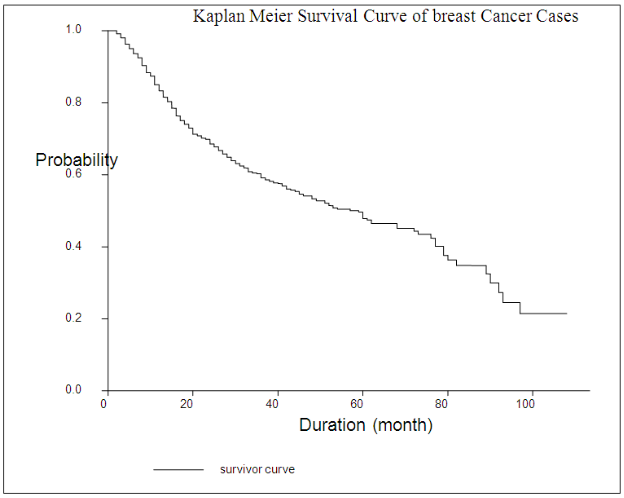

Fig.1. Survival Analysis for Breast Cancer Cases

The patients were divided into two groups; 49 years and below and greater than 49 years. Five hundred and eighty one of 1022 (56.85%) patients were 49 years of age and below, while the remaining were greater than 49years of age. In the survival function graph obtained by Kaplan – Meier which is a non-parametric technique, it is seen that the estimated survival probability of the 49 years and below group was slightly higher than the other group. The result of the log rank test used in comparison of the survival curves belonging to different age groups revealed that there was statistically no significant difference between the survival curves of the breast cancer patients ($\square^2$ = 0.715, p = 0.398).

Table 2: Log likelihood and AIC values of the Parametric models

Parametric model	Log-likelihood	AIC
Lognormal	-393.88	801.76
Weibull	-394.47	804.94
Gamma	-385.78	789.55
Exponential	-408.87	831.73
Log-logistics	-394.13	804.27

From table 2, the log likelihood and AIC statistics for the five parametric models applied to the survival data in order to select the best fitting parametric model. The parametric model was applied to the whole data set but not by age groupings since of no statistical significance between the groups. According to these criteria, the gamma model achieved the lowest AIC value of 789.55 and was therefore the best model for predicting the survival of breast cancer patients.

When analyzing the survival periods using parametric models, the size of tumour, tumour grade; stage; axillary node; BMI and Age were taken as the covariates. The hazard ratio obtained as a result of the analysis, p-value and 95% confidence interval are presented in table 3 using the gamma model. There isan increased risk of death in breast cancer cases: tumour size greater than or equal to 5cm compared with tumour size less than 5cm as the reference category (p= 0.0163);advanced tumour stage and not being obsessed. These factors were statistically significant with reduced survival of the breast cancer patients. The axillary nodes, age and tumour grade were not statistically significantly associated with increased risk of death following diagnosis of breast cancer.

Table 3: Covariates of survival of breast cancer cases diagnosed, applying the Gamma model

Covariates	Hazard Ratio	p-value	95% CI
Age	1.003	0.7862	0.987 – 1.01
Tumour size <5cm	1.611	0.0163	0.984 – 1.758
>5cm	0.000		
Axillary Node: <25%	0.976	0.7075	0.956 – 1.352
>25%	0.0000		
Tumour Grade: 1	0.953	0.3949	0.948 – 2.266
2	1.368	0.1103	0.936 – 1.92
3	0.000		
Tumour Stage: 0	1.470	0.0473	0.953 – 1.79
1	0.000		
BMI : 0	0.626	0.023	0.555 – 1.044
1	0.000		

IV. Discussion

Survival following a diagnosis of breast cancer was modeled by applying parametric survival models on the breast cancer patients. Applying likelihood based criteria for model selection indicated that the gamma model was the best fitting parametric model for predicting survival following breast cancer diagnosis. The mean age from the study was 47.97 years for the women diagnosed with breast cancer consistent with (Anyanwu, 2008), The age range was 20 – 92, with the highest incidence rate occurring in age group 40 - 49 years. This indicates that, Ghanaian women present breast cancer at a significantly younger age as compared to women in the developed world, (Adebamowo and Ajayi 2000; Schatzkin *et al.*, 1987; Muyberry and Stoddard-Wright 1992; Bowen *et al.*, 2008). The older one gets, the higher the risk of being diagnosed of breast cancer. Data relating to the clinical stages of breast cancer on first diagnosis showed that in all 52.35% of the women were presented at the advanced stages (III and IV) whiles early stage presentation involved 47.65%. This is consistent with some studies done in Ghana, (Clegg-Lamptey and Hodasi, 2007; Asumanu *et al.*, 2000 and Anim, 1979) and in Africa,(Boder, 2011; Ikpatt *et al.*, 2002). Differences among tumour grading were not significant as the p value was 0.3078. There was evidence of survivorship with BMI more than 25, (P = 0.013). the probability of surviving for obese women suffering from breast cancer has 37.9% greater than that of non-obese patients. The gamma model predicted a better survival for patients with smaller tumor size (< 5cm); early stage of diagnosis and obesity.

V. Conclusion

In summary, the study aimed to compare the results of the survival analysis of breast cancer patients using exponential, weibull, log-normal, gamma and log-logistic models. The data which comprised all Ghanaian women who reported at the Breast Clinic of the Korle_bu Teaching Hospital (KBTH) or its cancer centre between January 2002 and December 2008 were used. As a result of the analysis depending on the parameter estimates, women with bigger tumour size faced a higher risk of dying than the risk faced by those with smaller tumour size. Regarding BMI the claim that obese women faced a higher probability of surviving when compared to women with normal bodyweight was supported. Early staged breast cancer patients suffer a lower risk of death than the risk suffered by advanced staged breast cancers. Axillary node, tumour grade and age of the patient were not found as risk factors. By using AIC, the models obtained via exponential, weibull, log-normal, gamma and log-logistic were compared and the most suitable model for the obtained data distribution was the gamma distribution model. This is because it had the lowest AIC value as compared to the other models. Thus in conclusion the gamma distribution model is more suitable for these survival data obtained from KBTH.

Contribution to Knowledge

To the best of the researchers knowledge, this is the first study that survival following diagnosis of breast cancer is modeled using parametric models in Ghana. The study further examined the influence of some prognostic factors on breast cancer patients' survival. This study will contribute to the current body of knowledge on breast cancer survivorship and the factors that may improve a patient's quality of life during the transition period from being a patient to being a survivor.

Acknowledgment

We are thankful to the radiotherapy staff of the Korle - bu Teaching Hospital, who have kindly provided me with all the materials that I required for this study and for giving me the permission to browse through the files for the collection of my data.

References

[1] Adebamowo CA and Ajayi OO.2000. Breast Cancer in Nigeria.*West Afr J Med. Jul-Sep;19(3):179-91.*

[2] Allison, P. A. 1995. Survival analysis using the SAS system: a practical guide. SAS Institute, Inc., Cary, N.C.

[3] Amir H, Azizi MR, Makwaya CK, Jessani S., 1997. TNM classification and breast cancer in an African population: a descriptive study. Cent Afr J Med. ;43:357–359

[4] Amir H, Kitinya JN, Parkin BM. 1994. A comparative study of carcinoma of the breast in an African population. East Afr Med J. ;71:215–218.

[5] Anim JT.1979. Breast Cancer in Accra. *Ghana Med J*.;18:161–167.

[6] Anyanwu SNC. 2008. Temporal trends in breast cancer presentation in the third world. *J Exp Clin CancerRes.; 27(1): 17*

[7] Asumanu E, Vowotor R, Naaeder SB. 2000. Pattern of breast diseases in Ghana. *Ghana*

[8] *Med J.;34:206–209.*

[9] Ata N, Karasoy D, Sözer MT., 2008. Parametric and nonparametric survival models for nonproportional hazards. İstatistikciler Dergisi ;1(3):125-34.

[10] Boder J.M.E., F. B. E. Abdalla, M. A. Elfageih, A. Abusaa,A. Buhmeida, and Y. Collan,2011 "Breast cancer patients in Libya: comparison with European and central African patients,"*Oncology Letters*, vol. 2, no. 2, pp. 323–330.

[11] Bowen R.L. , S. W. Duffy, D. A. Ryan, I. R. Hart, and J. L. Jones, 2008 "Early onset of breast cancer in a group of British black women," *British Journal of Cancer*, vol. 98, no. 2, pp. 277–281.

[12] Bray F., MacCarron P. and D Maxwell Parkin, 2004 :The changing global patterns of female

[13] breast cancer incidence and mortality,Breast cancer Res,6:229-239

[14] Clegg-Lamptey JNA, Hodasi WH., 2007. A study of breast cancer in Korle Bu teaching hospital: Assessing the impact of health education. Ghana Med J.;41(2):72–77

[15] Gakwaya A, Kigula-Mugambe JB, Kavuma A, Luwaga A, Fualall J, Jombwe J, Galukande M, Kanyike D., 2008. Cancer of the breast: 5-year survival in a tertiary hospital in Uganda. Br J Cancer.;99:63–67.

[16] Gukas ID, Jennings BA, Mandong BM, Igun GO, Manasseh AN, Ugwu BT, Leinster SJ. 2005. Clinicopathololological features and molecular markers of breast cancer in Jos Nigeria. West Afr J Med. ;24:209–213.

[17] Hassan I, Onukak EE, Mabogunge OA., 1992. Breast cancer in Zaria, Nigeria. J R Coll Surg Edinb. 1992;37:159–161.

[18] IARC. International Association for Research on Cancer, *Handbook of Cancer 7th Ed. Lyon Press*. 2002:7-12.

[19] Ikpatt OF, Kuopio T, Collan Y. 2002. Proliferation in African breast cancer: biology and prognostication in Nigerian breast cancer material. *Mod Pathol* ;15:783–789.

[20] Mbonde MP, Amir H, Mbembati NA, Holland R, Schwartz-Albiez R, Kitinya JN. 1998. Characterisation of benign lesions and carcinomas of the female breast in a sub-Saharan population. Pathol Res Pract.;194:623–629

[21] Muyberry RM, Stoddard-Wright C., 1992. Breast cancer risk factors among black and white women: similarities and differences. *AmJEpidemioll* ;136:1445-1456.

[22] Okobia Michael N, Clareann H Bunker,Friday E Okonofua and Usifo Osime, 2006 :Knowledge attitude and practice of Nigerian women towards breast cancer : A cross sectional Study.World Journal of Surgical Oncology ;4 :11

[23] Schatzkin A, Palmer JR, Rosenberg L, et al.,1987. Risk factors for breast cancer in black women. *J Natl Cancer Inst*;78:213 – 7.

[24] Smith Robert A, Caleffi M., Ute-Susann Albert, Tony H H Chen, Stegen W. duffy and et al 2006: Breast cancer in limited resource countries: Early detection and access to care. The breast Journal,12(suppl.):S16-S26

[25] Wiredu EK, Armah HB., 2006. Cancer mortality patterns in Ghana: a 10 -year review of autopsies and hospital mortality. BMC Public Health. ;6:159–165.