Archives of Oncology and Cancer Therapy

ISSN: 2638-5074

Volume 2, Issue 2, 2019, PP: 26-39

Determinants of Survival Time of Women with Breast Cancer

M.K. Sharma, Endalkachew Abebe

Department of Statistics, Addis Ababa University, Addis Ababa, Ethiopia.
mk_subash@yahoo.co.in

***Corresponding Author:** M.K. Sharma, Department of Statistics, Addis Ababa University, Addis Ababa, Ethiopia.*

Abstract

Breast cancer is a cancer that develops from breast tissue and the most commonly diagnosed cancers worldwide with high death rate in the world. The aim of the study was to identify factors affecting the survival time of women with breast cancer. Total 819 women with breast cancer were included in the study based on the data taken from medical record card of patients enrolled during September 2016 to 2018 in Black Lion Hospital, Ethiopia. Kaplan-Meier plot (s) and Log-rank test were used for comparison of survival function; Cox-PH model and Accelerated failure time model were used to identify the factors which affect the survival time of women with breast cancer. Out of the total women 39.60% diedduring study. The median survival time of women was 33 months. The result of Weibull accelerated failure time (AFT) model showed that the survival time of woman with breast cancer significantly affected by age, oral contraceptives, alcohol consumption, breast feeding, tumor size, histologic grade and stages of breast cancer. The Weibull AFT model fit the breast cancer dataset better than other AFT models used in this study.

Keywords: *AFT, Weibull, women breast cancer, survival time*

Introduction

Cancer is a group of diseases characterized by the uncontrolled growth and spread of abnormal cells. If the spread is not controlled, it can result in death (Mandal et al., 2014). One of the most commonly diagnosed cancers worldwide was breast cancer, which accounts 1.7 million of the total, 521,900 deaths (IARC, 2013).

Breast cancer is a cancer that develops from breast tissue and most common invasive cancer in women (Nelson et al, 2013). In most countries, breast cancer is among the main causes of death in women (Fitzmaurice et al., 2015). According to GLOBOCAN 2018, an estimated 24.2% new breast cancer cases and 15.0% breast cancer-related deaths of women occurred worldwide.

Several studies has been done in survival data analysis using parametric and Semi-parametric survival models (Alizadeh et al., 2013; Hashemian et al., 2013; Vallinayagam et al., 2014; Pourhoseingholi et al., 2011 & Zare et al., 2013). As results from such studies shows the parametric survival models performing better than Semi-parametric model.

The Weibull model is most favorable for survival data analysis among different AFT models (Klein & Moeschberger, 1997). This was justified by study done using data on breast cancer cases by Ahmad et al. (2015) and Baghestani et al. (2015), which indicates that the Weibull model was chosen over different parametric models.

A study by Hoang (2014) on survival analysis of breast cancer patients, using Weibull model, revealed that age, stage of cancer, treatment by only surgery, treatment by both surgery and radiation has effect on patients' survival time.

In this article we did survival analysis using Cox-PH model to evaluate simultaneously the effect of several factorson women with breast cancer who receive treatments from September 2016 to 2018 inBlack Lion Hospital, Addis Ababa, Ethiopia. The time to death of women due to breast cancer was considered as the endpoint. In Cox-PH model, the proportional

hazards assumption holds with time-fixed covariates andcannot specify the general shape of the hazard curve (Collett, 2003). The AFT model is alternative over this model (Pourhoseingholi et al., 2011). In AFT models the effect size, is time ratio, is easier to interpret and more relevant to clinician than CoxPH model in which the effect size, is hazard ratio (Jiezhi, 2009).

Survival Model Used

Survival Function

The survivor function is the probability that the survival time of a randomly selected subject is greater than or equal to some specified time. Thus, it gives the probability that an individual surviving beyond a specified time. The distribution of survival time is characterized by survivorship, probability density and hazard function.

Let T be a random variable associated with the survival times and t be the specified value of the random variable T and f (t) be the underlying probability density function of the survival time T.

The survivor function S (t) is given by

$$S(t) = P(T > t) = 1 - F(t),\ t \geq 0$$

where, F(t) is cumulative distribution function, which represents the probability that a subject selected at random will have a survival time less than or equal to some stated value t, given by:-

$$F(t) = P(T \leq t) = \int_0^t f(u)du, t \geq 0$$

The probability density function, f (t), is given by:-

$$f(t) = \frac{d}{dt}F(t) = \frac{-d}{dt}S(t)$$

The hazard function is the instantaneous probability of having an event at time t (per unit time) given that one has survived (i.e. Not had an event) up to time t (Kleinbaum and Klein, 2011). It is given by:

$$\lambda(t) = \frac{f(t)}{S(t)}F(t) = \frac{-d}{dt}lnS(t)$$

The cumulative hazard function is defined as:-

$$\Lambda(t) = \int_0^t \lambda(u)du = -lnS(t)$$

Thus;

$$S(t) = e^{-\Lambda(t)}$$

Estimation of Survival Function

The Kaplan-Meier estimator is a non-parametric estimator, used to estimate the survival function with censoring on the order in which events and censored observations occur. Kaplan and Meier (1958) proposed the so called product –limit estimate of survival function:

$$\hat{S}(t) = \prod_{j:\tau_j \leq t} [1 - \frac{d_j}{r_j}]$$

where

- $\tau_j$denote the set of k distinct death time in the observed in the sample.
- $r_j$is the number of subjects alive (at risk) just before time t_j(the j^{th}ordered survival time).
- $d_j$denotes the number who died at time t_j.

Comparison of Survival Function

The Kaplan-Meier plots are used to see whether there is difference in survival time or not between groups of covariates under investigation. But the plot cannot be used to decide whether the survival time of women breast cancer in each covariate is different or not and for this purpose log-rank test is used (Mantel and Haenszel, 1959). The test statistic for log rank test is given by:

$$\chi^2_{logrank} = \frac{[\sum(d_{0j} - r_{0j}\frac{d_j}{r_j})]^2}{\sum \frac{r_{0j}r_{1j}d_j(r_j - d_j)}{r_j^2(1-r_j)}} \sim \chi^2_{(1)}$$

where

- $d_{0j}$is the number of failure in j^{th}time of 1^{st}group
- $d_{1j}$is the number of failure in j^{th}time of 2^{nd}group
- $d_j$is the number of failure in j^{th}time ($d_{0j} + d_{1j}$)
- $r_{0j}$is the number at risk at j^{th}time of 1^{st}group
- $r_{1j}$is the number at risk at j^{th}time of 2^{nd}group
- $r_j$is the number at risk at j^{th}time ($r_{0j} + r_{1j}$)

The hypotheses to be tested are:-

H_0: – There is no difference between the survival curves.

H_1: – There is difference between the survival curves.

Cox Proportional Hazards Model

The purpose of Cox PH model allows us to examine how specified factors influence the rate of a particular event happening at a particular point in time. This rate is commonly referred as the hazard rate. Thus, the relationship of predictors and the time-to-event in survival analysis is given through hazard function as follows:

$$\lambda(t|Z) = \lambda 0(t)e\beta 0Z = \lambda 0(t)e\beta 1Z1 + \ldots\ldots + \beta pZp$$

where

- $\lambda(t|Z)$ is the hazard at time t for a subject with a set of predictors $Z_1,...,Z_p$
- $\lambda_0(t)$ is the baseline hazard function, and
- $\beta_1,...,\beta_p$ are the model parameters describing the effect of the predictors on the overall hazard

The interpretation of the Cox PH model can be done using hazards ratios, which define as the ratio of two individuals with different covariate.

The corresponding survival function for Cox PH model is given by :

$$S(t|Z) = [S0(t)]e\text{^}\beta 1Z1+,...,+\beta pZp \quad (1)$$

where, $S_0(t)$ is the base line survival function

Accelerated Failure Time (AFT) Model

In most cases, the AFT models can provide more accurate estimates than CoxPH model. The key differences between the two models are baseline hazard function and ways of estimating coefficients (Kleinbaum& Klein, 2011).

The AFT model is parametric model and an alternative model to fit the survival data (Pourhoseingholi et al., 2011). Some of the standard parametric AFT models are Exponential, Weibull, Lognormal, and log logistic (Datwyler and Stucki, 2011).

The survival function of an individual with covariate X at time t, in the accelerated failure time models, is the same as the baseline survival function at time t is

$Exp(\beta_1X_{1i}+,.....,+\beta_pX_{pi})$, where $\beta_1,...,\beta_p$ are coefficients of the regression. Thus the survival function of time t,

$$S(t|X) = S_0[t * exp(\beta_1X_{1i}+,.....,+\beta_pX_{pi})] \text{ for all } t\geq 0.$$

The effect of the covariates on the survival function is changed by a factor $exp(\beta^0X)$ which is called accelerated factor.

The AFT model treats the logarithm of survival time as the response variable and includes an error term that is assumed to follow a particular distribution. The AFT model can be written as follows:-

$$logT_i = \mu + \beta_1X_{1i}+,.....,+\beta_pX_{pi}+ \sigma\varepsilon_i$$

where: μ is an intercept, $logT_i$is the log-transformed survival time, $X_1,...,X_p$are explanatory variables with coefficients $\beta_1,...,\beta_p$, $\varepsilon_i$represents residual or unexplained variation in the log-transformed survival times and σ is the scale parameter.

Table 2.1. *Commonly used distributions and parameters in AFT models*

Distribution	f(t)	S(t)	λ(t)
Exponential	$\Lambda e-\lambda t$	$e\text{^}-\lambda t$	λ
Weibull	$\lambda\rho tp-1e-\rho t$	$e-\rho t$	$\lambda\rho tp-1$
Log-logistic	$\frac{\lambda\rho t^{\rho-1}}{[1+\lambda\rho t^{\rho}]^2}$	$\frac{1}{1+\lambda\rho t^{\rho}}$	$\frac{\lambda\rho t^{\rho-1}}{1+\lambda\rho t^{\rho}}$
Log-normal	$\frac{1}{\sqrt{2\pi}\sigma}exp[\frac{-[logt-\mu]^2}{2\sigma^2}]$	$1-\Phi[\frac{logt-\mu}{\sigma}]$	$\frac{\frac{1}{\sqrt{2\pi}\sigma}exp[\frac{-[logt-\mu]^2}{2\sigma^2}]}{1-\Phi[\frac{logt-\mu}{\sigma}]}$

where

- λ and ρ denotes scale parameter and shape parameter respectively for Exponential, Weibull, & Log-logistic distribution.
- σ and μ denote scale parameter and shape parameter respectively for Log-normal distribution.
- $\Phi(.)$ denotes the standard normal distribution function.

Methods of Statistical Analysis

Descriptive Statistics

The frequency distribution table was also used to summarize the data obtained from registration book of patients based on the study variables in Oncology Department of Black Lion Hospital, Addis Ababa, Ethiopia.

Estimation of Parameters in AFT Model

The parameters of Cox-PH model were estimated by partial likelihood function. (Cox, 1972).

The parameters of AFT models were estimated by maximum likelihood method and Newton-Raphson procedure (Jiezhi, 2009).

Model Building

We used Hosmer and Lemeshow (1998) recommendations in selecting studyvariables.

Model Selection Criterion

The Akaike's Information Criterion (AIC) was used to choose the best AFT model from Exponential, Weibull, Log-logistic and Log-normal model. (Akaikie, 1974).

The model with smaller AIC is better as compared to other. AIC is obtained by:-

$$AIC = -2log\,(L) + kp$$

where:

- p is the number of parameters in the model
- L is the likelihood
- k is a constant and can be seen as a penalty for additional parameters between 2 and 6 (often 2). The recommendation is to use a larger k with small sample.

Model Diagnostics

Checking the Assumption of Cox PH Model

The assumptions of Cox-PH model was checked by test of correlation (rho) and global test. The assumptionsare valid if the test result is insignificant.

Adequacy Checking for the Parametric Baselines Distribution

The plot results (Datwyler and Stucki 2011)

- $log[S(t)]$ Vs t gives approximately a straight line that pass through the origin if exponential distribution is reasonable.
- $log[-log[S(t)]]$ Vs $log[t]$ gives approximately a straight line (linear) if weibull distribution is reasonable.
- $log[S(t)/1 - S(t)]$ Vs $log(t)$ is linear if log-logistic distribution is reasonable.
- $\Phi^{-1}[1-S(t)]$ Vs $log(t)$ is linear if log-normal distribution is reasonable. where:

- $S(t)$ denotes the corresponding survival function of baseline distribution and estimated using kaplan-Meier estimate.
- $\Phi(.)$ denotes the standard normal distribution function.

Residual Plot

The Cox-Snell residuals, r_j, is defined by:

$$r_j = \Lambda(\hat{T}_j|X_j)$$

Where $\Lambda\hat{}$ is the cumulative hazard function of the fitted model. If the model fits the data, then the r_j's should have a standard ($\lambda = 1$) exponential distribution, so that a plot of r_j versus the Nelson-Aalen estimator of the cumulative hazard of the r_j's should be approximately straight line with slope 1 (Jiezhi, 2009).

Data

Study Area

The study has been conducted on the data taken from Black Lion Hospital, Ethiopia. The cancer unit at the Black Lion Hospital provides chemotherapy, radiotherapy; hormone therapy and other supportive and palliative cares. It is the main center for cancer registry, early detection, prevention, standard treatment and palliative care in Addis Ababa. It started an organized oncology service in 1998 Ethiopian calendar.

Study Population

A retrospective study has been conducted on the data taken on breast cancer of women recorded in Oncology Department of Black Lion Hospital, Addis Ababa, Ethiopia. The population of this study was all women with breast cancer who had been registered at Oncology Department of Black Lion Hospital for three years starting from September, 2016 up to 2018. The data has been carefully reviewed from the registration log book and patients' registration card; any inadequate information encountered was checked from the file and excluded from analysis if proven to be inadequate. Thus, the data has been collected from patient follow up records based on the variables in the study.

Inclusion and Exclusion Criteria

Inclusion Criteria: All women with breast cancer registered with full information including study variables of interest in the registration book or in the chart were considered to be eligible for the study.

Exclusion Criteria: Women with insufficient information regarding study variables on the registration book or in the card were not eligible.

Data Collection Procedure

Ethical permission has been obtained from the Oncology Department of Black Lion Hospital, Addis Ababa, Ethiopia. Then secondary data were taken from existing records in the hospital by trained enumerator and theresearcher.

Variables in the Study

The response variable was survival time of women

with breast cancer (in months), defined as the difference between time of diagnosis and time to one of the events "death", "lost to follow up", "dropped out", "stopped", "transferred out to other health centers or hospitals" occurred. Death was considered to be the event of interest.

The factors considered for the purpose of survival analysis of women with breast cancer were as follows:-

1. Age (≤ 30, 31-49, ≥ 50)
2. Region (Addis Ababa, Oromia, Amhara, Southern Nation and Nationality Peoples, Others)
3. Residence (Urban, Rural)
4. Oral contraceptives (Not used, used)
5. Alcohol consumption (No, Yes)
6. Family history of breast cancer (No, With)
7. Breast Feeding (No, Yes)
8. Recurrence (No, Yes)
9. Tumor size (≤ 2 cm, 2-5 cm , >5cm)
10. Histologic grade (I, II, III)
11. Treatments taken (Chemotherapy, Radiotherapy, Surgery, Hormone Therapy, Combination of the two or more)
12. Stages of breast cancer (I, II, III, IV)

Results

Descriptive Analysis

The minimum and maximum event time observed for women with breast cancer follow up were 6 and 36 months, respectively. Among those women with breast cancer patients, about 60.40% were censored and the remaining 39.60% were died. Fifty percent of women with breast cancer patients survived 33 months or above it.

In Appendix Table 5.1, exhibits, the survival time of women with breast cancer inAddis Ababa, Oromia, Amhara, and SNNP as well as in all other region 38.10%, 27.40%, 15.50 %, and 9.50%, respectively. The survival time of women appears higher than women using oral contraceptive. The tumor size of 27.4%, 34.10% and 38.60% of women with breast cancer have a size of ≤ 2 cm, 2-5 cm and >5 cm, respectively. The women having higher tumor size have low survival time. The women, who go for treatment into the hospital at the later stage, have low survival time.

Comparison of Survival Time of Women Patients

The log-rank test in Table 5.2 shows, there is significant difference in the survival time of women with breast cancer for age, oral contraceptives, alcohol consumption, family history of breast cancer, breast feeding, recurrence, tumor size, histologic grade, treatments taken and stages of breast cancer since there corresponding p-values are smaller than the common level of significance (5 %).

Table 5.2. *Log-rank test of each the covariates*

Covariates	Chi-square	Degree of freedom	P-value
Age	19.80	2	<0.0001
Residence	0.60	1	0.5000
Breast Feeding	5.60	1	0.0200
Oral Contraceptives	4.00	1	0.0470
Alcohol Consumption	12.10	1	0.0005
Family History of Breast Cancer	5.10	1	0.0200
Recurrence	7.40	1	0.0070
Tumor Size	9.10	2	0.0100
Histologic Grade	6.10	2	0.0490
Treatments Taken	12.00	4	0.0200
Stages of Breast Cancer	139.00	3	<0.0001

Cox Proportional Hazard Model

Checking the Assumption of Cox-PH

Table 5.3. *Shows test of assumption in Cox-PH model*

Covariates	rho	Chi-square	P-value
Age	0.00791	0.0205	0.8863
Residence	0.60	0.1216	0.5000
Breast Feeding	0.07917	2.1373	0.0148
Oral Contraceptives	-0.01793	0.1216	0.1273
Alcohol Consumption	0.02202	0.1829	0.6689
Family History of Breast Cancer	-0.04066	0.4827	0.4872
Recurrence	0.04570	0.6259	0.1429
Tumor Size	0.06610	1.4410	0.0300
Histologic Grade	0.07225	1.7019	0.1920
Treatments Taken	0.12584	5.0502	0.0185
GLOBAL TEST	NA	11.2956	0.0260

The values of rho in Table 5.3 above exhibits that the assumptions of Cox-PH model completely failed because the values for breast feeding, tumor size and treatments (taken) are less than (5%) level of significance. For global test the assumptions of Cox-PH also failed due to significance of the result.

Accelerated Failure Time Models

For the survival time of women with breast cancer data, we fitted AFT models such as Exponential, Weibull, Log-logistic and Lognormal distribution taking all the covariates found to be significant in the uni-variable case at 5% level of significance. The summary of uni-variable AFT model analysis is given in Table 5.4.1 Appendix.According to the Table, the AFT (Exponential, Weibull, Log-logistic & Log-normal) models shows that age, alcohol consumption, family history of breast cancer, breast feeding, recurrence, tumor size, histologic grade, treatments (taken) and stages of breast cancer are significantly associated with survival time of women with breast cancer at 5% level of significance. In Weibull distribution oral contraceptives was also significantly associated with survival time of women at 5% level of significance.

To compare the efficiency of these models AIC is used and the one with the smallest AIC fit the data very well than the others. All AFT models and the corresponding AIC values are given in Table 5.4.2 below. Weibull AFT model (AIC =2754.316) found to be the best for the survival time of women with breast cancer.

Table 5.4.2. *Comparisons of AFT models using AIC*

Distribution	AIC
Exponential	3146.397
Weibull	**2754.316**
Log-Logistic	2774.607
Log-normal	2795.853

Weibull Accelerated Failure Time Model

The final results for the Weibull AFT model is shown as in Table 5.4 and we find that the survival time of woman with breast cancer significantly affected by age, oral contraceptives, alcohol consumption, breast feeding, tumor size, histologic grade, treatments taken and stages of breast cancer.

Table 5.5. *Result of maximum likelihood parameter estimates of the Weibull AFT model*

Covariates	Categories	$\hat{\beta}$	SE[$\hat{\beta}$]	φ	[95%$CI\hat{\varphi}$]	p-value
Intercept	≤ 30	4.3987	0.1499			
Age	31-49	-0.1537	0.0676	0.858	[0.751 , 0.979]	0.0231
	≥ 50	-0.1942	0.0680	0.823	[0.721 , 0.941]	0.0043
Breast feeding	No					
	Yes	0.0976	0.0374	1.103	[1.025 , 1.186]	0.0091
Oral contraceptives	Not used					
	Used	-0.2596	0.0818	0.771	[0.804 , 0.992]	0.0015
Alcohol consumption	No					
	Yes	-0.3462	0.0828	0.707	[0.601 , 0.832]	<0.0001
	≤ 2 cm					
Tumor size	2-5 cm	-0.0893	0.0501	0.915	[0.829 , 0.998]	0.0145
	>5cm	-0.0990	0.0480	0.906	[0.824 , 0.995]	0.0392
	I					
Histologic grade	II	0.0466	0.0430	1.048	[0.963 , 1.140]	0.2777
	III	0.1435	0.0475	1.154	[1.0517 , 1.267]	0.0025
	I					
Stages of breast cancer	II	-0.5598	0.1320	0.571	[0.441 , 0.740]	<0.0001
	III	-0.7586	0.1308	0.468	[0.362 , 0.605]	<0.0001
	IV	-0.9251	0.1313	0.397	[0.307 , 0.513]	<0.0001

φ: indicates Acceleration factor; 95%CI for φ: 95% confidence interval for acceleration factor; SE: standard error

Interpretation of Weibull AFT Model

Under the Weibull AFT model, keeping the effect of other factors constant, the estimated acceleration factor for women with breast cancer patients of age 31 to 49 and 50 or above 50 years old are estimated to be 0.858 with [95 % CI: 0.751 , 0.979] and 0.823 with [95% CI: 0.724 , 0.941] respectively. Thus, the expected survival time of women with breast cancer patients decrease by 14.20% and 17.70% for women aged 31 to 49 and 50 or above 50 years older, respectively, as compared to women with breast cancer patients of age 30 or below 30 years. The 95% confidence interval for acceleration factors of both age group did not include one and p-values are small (p-value=0.0231 and p-value= 0.0043) which implies that both age group has significant effect on the survival time of women with breast cancer patients.

By observing breast feeding of women with breast cancer patients, keeping the effect of other factors constant, the estimated acceleration factor for women in breastfed a child is estimated to be 1.103 with [95% CI: 1.025 , 1.186] in which the expected survival time is 10.30% higher than women did not breastfed child. The 95% confidence interval for acceleration factor of women breastfed a child did not include one and p-value is small (0.0091) which implies that women breast feeding has significant effect on the survival time of women with breast cancer .

Looking for women oral contraceptive use, keeping the effect of other factors constant, the estimated acceleration factor of women using oral contraceptive is estimated to be 0.771 with [95% CI: 0.804 , 0.992] which implies the expected survival time decreases by 22.90% than women did not use oral contraceptive. The 95% confidence interval for acceleration factor of women oral contraceptive use did not include one and p-value is small (0.0015) which implies that women oral contraceptive has significant effect on the survival time of women with breast cancer patients.

On other hand, keeping the effect of other factors constant, the estimated acceleration factor for women patients consuming alcohol is estimated to be 0.707 with [95% CI: 0.601, 0.832]. The 95% confidence interval for the acceleration factor did not include

one and p-value is small (P-value=<0.0001). Thus, women consuming alcohol has significant effect on the survival time of patients and the expected survival time of women with breast cancer consuming alcohol decreases by 29.30% than not consuming alcohol.

Regarding tumor size of women breast cancer, keeping the effect of other factors constant, the estimated acceleration factor for women with breast cancer patients of tumor size 2 to 5 and above 5 centimeters are estimated to be 0.915 with [95% CI: 0.829 , 0.998] and 0.906 with [95% CI: 0.814 , 0.985] respectively. Thus, the expected survival time of women with breast cancer decreases by 8.50% for tumor size 2 to 5 centimeters and 9.40% for tumor size above 5 centimeters of women with breast cancer as compared to women with breast cancer patients of tumor size 2 or below 2 centimeters. The 95% confidence interval for acceleration factor of both tumor size of women breast cancer did not include one and p-value are small (p-value=0.0145 and p-value=0.0392) which implies that both tumor size of women with breast cancer has significant effect on the survival time of women with breast cancer patients.

Moreover, for those patients having histologic grade three, keeping the effect of other factors constant, the estimated acceleration factor is estimated to be 1.154 with [95% CI: 1.052 , 1.267] which implies that the expected survival time increases by 15.40% than women having histologic grade one. The 95% confidence interval for acceleration factor of women breastfed a child patients did not include one and p-value is small (0.0025) which implies that women having histologic grade three has significant effect on the survival time of women with breast cancer patients.

Finally, observing for stages of women with breast cancer, keeping the effect of other factors constant, the estimated acceleration factor for stage II, III and IV of breast cancer are estimated to be 0.571 with [95% CI: 0.441 , 0.738], 0.468 with [95% CI: 0.362 , 0.605] and 0.397 with [95% CI: 0.307 , 0.513] respectively. Thus, the expected survival time of women with breast cancer decreases by 42.90%, 53.20% and 60.30% for stage II, III and IV of women with breast cancer respectively as compared to stage one. The 95% confidence interval for acceleration factor of those three stages of breast cancer did not include one and p-value are small (<0.0001) and this indicates that stage II, III and IV of breast cancer has significant effect on the survival time of women with breast cancer patients.

Model Diagnostics

Graphical Adequacy Checking of Parametric Baselines

It can be observed from Figure1, below, the plots for baseline distribution and the plot for Weibull distribution are approximately straight line or linear. Thus, the plot result shows fitting the Weibull distribution is reasonable.

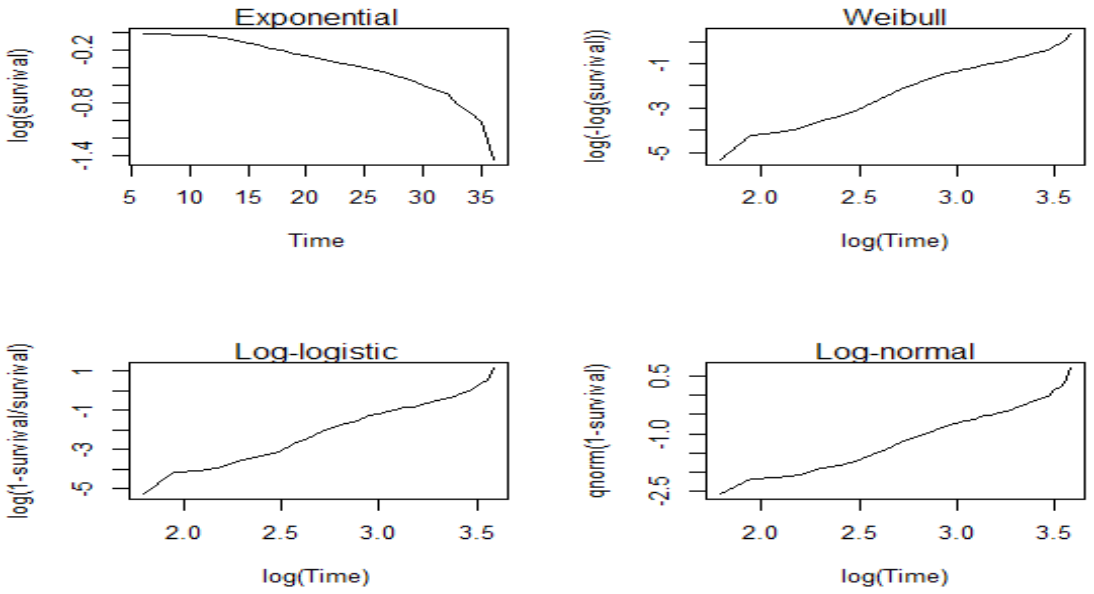

Figure 1. *Figure shows baseline distributions plot for survival time of women breast cancer patients' data*

Cox-Snell Residuals Plot

By observing Cox-Snell residuals plot Figure 2 below, the Weibull AFT model fits women with breast cancer patients data, since the plot of Cox-Snell residuals against cumulative hazard function of residuals is approximately a straight line with slope one.

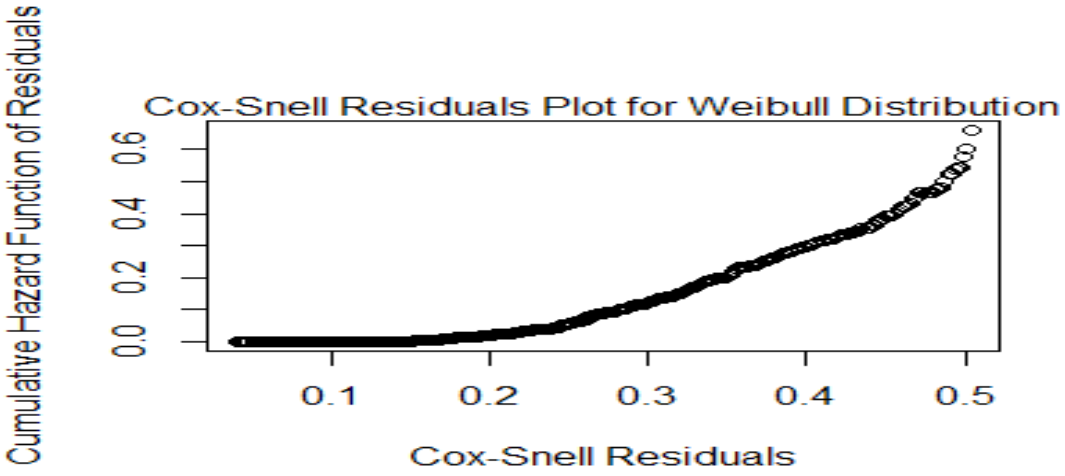

Figure 2. *Figure shows Cox- Snell residuals plots of Weibull distribution for survival time of women breast cancer patients' data*

Likelihood Ratio Significance Test of the Weibull AFT Model

The likelihood ratio test in Table 5.5 below shows that the Weibull AFT model is significant. In addition, the likelihood ratio test shows that the Weibull AFT model can fit women breast cancer data well since the likelihood values of the full model has improvement after the covariates were added in the model than likelihood of null model.

Table 5.5. *The likelihood ratio significance of Weibull AFT model*

Loglik(intercept only)	Loglik(model)	Chi-square(χ^2)	Degree of Freedom (Df)	p-value
-1463.00	-1361.20	203.450	12	<0.0001

Discussion

The main aim of this study was to identify factors affecting the survival time of women breast cancer data set, which was obtained from Black Lion Hospital. The Cox-PH model was applied for this data but its assumptions were violated. Therefore the AFT models with baseline distribution: Exponential, Weibull, Log-logistic and Log-normal were used.. To compare the efficiency of different AFT models AIC was calculated and Weibull AFT model found to be the best for the survival time of women with breast cancer. The result of this study was consistent with study done on breast cancer cases by Ahmad et al. (2015) and Baghestani et al. (2015).

Under uni-variable analysis the Weibull AFT model shows that age, alcohol consumption, family history of breast cancer, breast feeding, oral contraceptives, recurrence, tumor size, histologic grade, treatments taken and stages of breast cancer were significantly associated with survival time of women with breast cancer patients at 5% level of significance.

From result of multivariable analysis of Weibull AFT model the survival time of woman with breast cancer patients significantly affected by age, oral contraceptives, alcohol consumption, breast feeding, tumor size, histologic grade and stages of breast cancer. The results of this study was in agreement with a study by Hoang (2014) on survival time of breast cancer patients using Weibull AFT model which indicates that age and stages of breast cancer has effect on survival time of women with breast cancer patients.

From the result of this study the age of women breast cancer patients has a significant effect on the survival time of women breast cancer and their survival time

decreases as they gets older. The study by Hoang in 2014 also showed the same results.

The survival time of women with breast cancer patients significantly affected by women breast feeding and the survival time of women breastfed their child was higher than that of do not breastfed child.

On other hand, the survival time of women consuming alcohol decreases as compared to women did not consume alcohol and this result was consistent with study done by Chen et al. (2011).

The study done by Urban et al. (2012) showed that the survival time of women with breast cancer patients significantly affected by oral contraceptives and the expected survival time of women using oral contraceptives was less as compared to women did not use oral contraceptives. This result of our study was consistent with the current study.

The survival time of women with breast cancer of tumor size 2 to 5 centimeters and above 5 centimeters were decrease as compared to women with breast cancer of tumor size 2 or below 2 centimeters. This result was justified bythe study by Mensah et al. (2016) in Ghana.

In addition, the histologic grade at diagnosis was significantly affected the survival time of women with breast cancer and the survival time was high for women in histologic grade three as compared to women in histologic grade one. This was also reported by Ahmad et al in 2015.

The stages of breast cancer have significant effect on the survival time of women with breast cancer patients. The study done by Allemani et al. (2015) and Hoang (2014) also showed that the stages of breast cancer at diagnosis has been significantly affected the survival time of women with breast cancer patients. From the results of these two studies the survival time of women with breast cancer was smaller as the stage increases as in result of this study.

References

[1] Aalen, O. (2008) *Survival and Event History Analysis*. Springer-Verlag, New York.

[2] Abate SM., Yilma Z, Assefa M., & Tigeneh W. (2016) *Trends of Breast Cancer in Ethiopia.* Int J Cancer Res Mol Mech 2(1): doi http://dx.doi.org/10.16966/2381-3318.121

[3] Ahmad R.B., Sahar S.M., Hamid A.M., Mohammad E.A., Nahid N., & Kimiya G. (2015) *Survival Analysis of Breast Cancer Cases in Iran using a Weibull Parametric Model:* Asian Pac J Cancer Prev, 16 (18) , 8567-8571

[4] Akaike H. (1974) *A new look at the statistical model identification*. IEEE Trans. Automat. Contr. AC-19:716-23. [Institute of Statistical Mathematics, Minato-ku, Tokyo, Japan]

[5] Alizadeh A., Mohammadpour RA., & Barzegar MR. (2013) *Comparing cox model and parametric models in estimating the survival rate of patients with prostate cancer on radiation therapy*. Journal of Mazandaran University of Medical Sciences (JMUMS), 23.

[6] Allemani C., Weir HK., Carreira H., Harewood R, Spika D, Wang XS, et al. 'Global surveillance of cancer survival 1995-2009: analysis of individual data for 25,676,887 patients from 279 population-based registries in 67 countries (CONCORD-2). Lancet. 2015;385(9972):977-1010.

[7] Amend K., Hicks D., & Ambrosone CB. (2013) *Breast cancer in AfricanAmerican women: Differences in tumor biology from EuropeanAmerican women.* Cancer Res 66:83278330.

[8] American Cancer Society (2017) *Breast Cancer Facts & Figures 20172018*. Atlanta: Inc. 2017.

[9] American Cancer Society (2018) *Cancer Facts & Figures*. Atlanta: Inc.2018

[10] American Joint Committee on Cancer Breast Cancer Staging (2016) 7th Edition https://cancerstaging.org/referencestools/quickreferences/Documents/BreastMedium.pdf (accessed 22 November 2016).

[11] Baghestani A., Moghaddam S., Majd H., Akbari M, Nafissi N, Gohari K et al. (2015) *Survival Analysis of Patients with Breast Cancer using Weibull Parametric Model*. Asian Pac J Cancer Prev 2015, 16(18):85678571.

[12] Chen WY., Rosner B., Hankinson SE., Colditz GA., & Willett WC. 'Moderate alcohol consumption during adult life, drinking patterns, and breast cancer risk. JAMA. 2011;306: 1884-1890.

[13] Collett D. (2003) *Modeling survival data in medical research*. Chapman & Hall, London

[14] Cox D. (1972) *Regression models and life tables (with discussions)*. Journal of the Royal Statistical Society. 34: 187-220

[15] Dai X., Chen A., &Bai Z. (2014) 'Integrative investigation on breast cancer in ER, PR and HER2-defined subgroups using mRNA & miRNA expression profiling. Sci Rep 4: 6566.

[16] Datwyler C., & Stucki T. (2011) *Parametric Survival Model*: *https://stat.ethz.ch/education/semesters/ss*2011*/seminar/contents/handout*9*. pdf*

[17] Ferlay J., Soerjomataram I., & Ervik M. (2014) *GLOBOCAN 2012 v1. 0, Cancer Incidence and Mortality Worldwide*: IARC CancerBase No. 11. Lyon, France: International Agency for Research on Cancer; 2013. Visit: http://globocan. iarc. fr.

[18] Fitzmaurice C., Dicker D., & Pain A. (2015) 'The Global Burden of Cancer 2013. JAMA Oncol, 1, 505-27.

[19] Global Burden of Disease Cancer Collaboration (2015) *Global, Regional, and National Cancer Incidence, Mortality, Years of Life Lost, Years Lived With Disability, and Disability-Adjusted Life-years for 32 Cancer Groups, 1990 to 2015*: A Systematic Analysis for the Global Burden of Disease Study. JAMA Oncol. 2017 April 01; 3(4): 524548. doi:10.1001/jamaoncol.2016.5688.

[20] GLOBOCAN (2018) *Estimates of Incidence and Mortality Worldwide for 36 Cancers in 185 Countries*: doi: 10.3322/caac.21492. Available online at cacancerjournal.com

[21] Hashemian AH., Beiranvand B., & Rezaei M., (2013) A comparison between cox regression and parametric methods in analyzing kidney transplant survival. World Applied Sciences Journal, 26, 502-7.

[22] Hoang P. (2014) *Survival Analysis - Breast Cancer, Undergraduate*. Journal of Mathematical Modeling, DOI: http://dx.doi.org/10.5038/2326-3652.6.1.4860

[23] Hosmer D., & Lemeshow S., (1998) *Applied Survival Analysis Regression Modeling of Time to Event Data*. John Wiley and Sons, New York.

[24] International Agency for Research on Cancer (2013) *Latest world cancer statistics Global cancer burden in 2012*: Marked increase in breast cancers must be addressed. WHO: Press release N 223.

[25] Jiezhi Qi. (2009) *Comparison of Proportional Hazards and Accelerated Failure Time Models* https://core.ac.uk/download/pdf/55298480.pdf

[26] Kantelhardt EJ., Zerche P., Mathewos A., Trocchi P., & Addissie A. (2014) *Breast cancer survival in Ethiopia: A cohort study of 1,070 women*. Int J Cancer 135: 702-709.

[27] Kaplan EL. & Meier P. (1958) *Non-parametric estimation from incomplete observations*. J am Stat Assoc 1958(53):45781

[28] Klein D. G. (2005) *Survival Analysis, third edition*. Springer, New York

[29] Klein JP. & Moeschberger ML. (1997) Survival analysis techniques for censored and truncated data. Springer, New York

[30] Kleinbaum DG. & Klein M (2011) *Survival analysis a self-learning text, 3rd edn*. Springer, New York

[31] Kotepui M. & Chupeerach C. (2013) *Age distribution of breast cancer from a Thailand population- based cancer registry*. Asian Pac J Cancer Prev 14: 3815-3817.

[32] Mandal A., Islam NK., Scott J., Okafor B., & Mandal PK. (2014) *African Americans and Cancer: A Minority Health Advocacy*. Journal Bioproces Biotech 4: e117.

[33] Mantel N. & Haenszel W. (1959) *Statistical aspects of the analysis of data from retrospective studies of disease*. J Natl Cancer Inst 22:719748

[34] Mensah AC., Yarney J., Nokoe SK., Opoku S., & Clegg-Lamptey J. (2016) *Survival outcomes of breast cancer in Ghana: an analysis of clinicopathological features*. Open Access Library J 2016, 3:1-11.

[35] Nelson H.D., Smith M.E., Griffin J.C., & Fu R., (2013) 'Use of medications to reduce risk for primary breast cancer: a systematic review for the U.S. Preventive Services Task Force. Annals of Internal Medicine 158: 604-14.

[36] Nuno R., Coleman K., Bengoa R., & Sauto R. (2012) 'Integrated care for chronic conditions:

the contribution of the ICCC Framework. Health Policy 2012, 105(1):55-64.

[37] Parkin, D. M., F. Bray, J. Ferlay, & A. Jemal. (2014) 'Cancer in Africa 2012. Cancer Epidemiol Biomarkers Prev 23 (6):953-66. doi: 10.1158/1055-9965.epi-14-0281.

[38] Pourhoseingholi MA., Pourhoseingholi A., & Vahedi M (2011) *Alternative for the Cox regression model: using parametric models to analyze the Survival of Cancer Patients*. Iran J Cancer Prev; 4:1- 9.

[39] Rahimzadeh M., Baghestani AR., & Gohari MR., (2014) 'Estimation of the cure rate in Iranian breast cancer patients. Asian Pac J Cancer Prev,15, 4839-42.

[40] Stewart BW. & Wild CP. (2014) *World Cancer Report; 3rd edition*. IARC, France.

[41] Tigeneh W., Molla A., Abreha A., & Assefa M. (2015) *Pattern of Cancer in TikurAnbessa Specialized Hospital Oncology Center in Ethiopia from 1998 to 2010*: Int J Cancer Res MolMech 1.1.

[42] Urban M., Banks E., Egger S., Canfell K., & OConnell D. (2012) 'Injectable and oral contraceptive use and cancers of the breast, cervix, ovary, and endometrium in black South African women: casecontrol study. PLoS Med 9: e1001182.

[43] Vallinayagam V., Prathap S., & Venkatesan P. (2014) *Parametric regression models in the analysis of breast cancer survival data*. International Journal of Science and Technology, 3, 163-7.

[44] Woldeamanuel Y.W., Girma B., & Teklu A.M. (2013) 'Cancer in Ethiopia'. Lancet Oncol 14 (4):289-90. doi: 10.1016/s1470-2045(12)70399-6.

[45] Woldu MA., Legese DA., Abamecha FE., & Berha AB. (2017) *The Prevalence of Cancer and its Associated Risk Factors among Patients Visiting Oncology Unit, TikurAnbessa Specialized Hospital, Addis Ababa-Ethiopia*. J Cancer SciTher 9: 414-421. doi: 10.4172/19485956.1000452

[46] World Health Organization (2014) *Breast Cancer Facts and Figures from 2013 to 2014*. The documents available online at: http://www.cancer.org/acs/groups/content/@research/documents/document/acspc042725. Accessed on: 6/21/2015.

[47] World Health Organization (2015) *Cancer - a Growing Public Health Concern for Ethiopia*. available at www.who.int

[48] Yip CH., Buccimazza I., & Hartman M., (2015) *Improving outcomes in breast cancer for low and middle income countries*. World J Surg, 39, 686-92.

[49] Zare A, Mahmoodi M, & Mohammad K, (2013) Comparison betweenparametric and semi-parametric cox models in modeling transition rates of a multi-state model: application in patients with gastric cancer undergoing surgery at the Iran cancer institute. Asian Pac J Cancer Prev,14, 6751-5.

Appendix

Table 5.1. *Descriptive Statistics summary*

Patients Status

Covariates	Categories	N of Censored (%)	N of Death(%)	Total
	≤ 30	91(18.40)	29(9.00)	120(14.70)
Age	31-49	235(47.50)	144(44.40)	379(46.30)
	≥ 50	169(34.10)	151(46.60)	320(39.10)
	Addis Ababa	184(37.20)	128(39.50)	312(38.10)
	Oromia	140(28.30)	88(27.20)	228(27.40)
Region	Amhara	84(17.54)	43(13.30)	127(15.50)
	SNNP	44(8.90)	34(10.50)	78(9.50)
	Others	43(8.70)	31(9.60)	74(9.50)
Residence	Urban	262(52.90)	158(48.80)	420(51.30)

	Rural	233(47.10)	166(51.20)	399(48.70)
Breast feeding	No	228(46.10)	175(54.40)	403(48.20)
	Yes	267(53.90)	149(46.60)	416(50.80)
Oral contraceptives	Not used	262(52.90)	146(45.10)	408(49.80)
	used	233(47.10)	178(54.90)	411(50.20)
Alcohol consumption	No	258(52.10)	133(41.00)	391(47.70)
	Yes	237(47.90)	191(59.00)	428(52.30)
Family history of breast cancer	No	196(39.60)	161(49.70)	357(43.60)
	With	299(60.40)	163(50.30)	462(56.40)
Recurrence	No	187(37.80)	159(49.10)	346(42.20)
	Yes	308(62.20)	165(50.90)	473(57.80)

Extension of Table above

Patients Status

Covariates	Categories	N of Censored(%)	N of Death(%)	Total
Tumor size	≤ 2cm	152(30.70)	72(22.20)	224(27.40)
	2-5cm	164(33.10)	115(35.50)	279(34.10)
	>5	179(36.20)	137(42.30)	316(38.60)
Histologic grade	I	136(27.50)	108(39.50)	312(38.10)
	II	140(28.30)	88(27.20)	228(27.40)
	III	84(17.54)	43(13.30)	127(15.50)
Treatments taken	Chem	152(30.70)	102(31.50)	254(31.00)
	Rad	79(16.00)	36(11.10)	115(14.00)
	Sur	47(9.50)	31(9.60)	60(7.30)
	Hor	47(9.50)	47(9.50)	47(9.50)
	Comb>2	47(9.50)	47(9.50)	47(9.50)
Stage of breast cancer	I	129(26.10)	7(2.20)	136(16.60)
	II	162(32.70)	62(19.10)	124(27.40)
	III	130(26.30)	106(32.70)	236(28.80)
	IV	74(14.90)	149(46.00)	223(27.20)

where N denote number , SNNP is South Nation Nationality of Peoples and others stands for region not listed here (Tigray, Harare, Diredawa, Benishangul-gumuz, Gambela, Somalie, Afar, Non-identified)

Table shows Counts for patient status

Patient Status	Number of Patients (%)
Censoring	495(60.40)
Death	324(39.60)

Table 5.4.1. *Uni-variable AFT models analysis results using different baseline hazard function*

		Distributions			
		Exponential	Weibull	Log-logistic	Log-normal
Covariates	Categories	$\hat{\beta}$ [p-value]	$\hat{\beta}$ [p-value]	$\hat{\beta}$ [p-value]	$\hat{\beta}$ [p-value]
Age	≤ 30				
	31-49	-0.483[0.018]	-0.182[0.011]	-0.218 [0.003]	-0.254[0.001]
	≥ 50	-0.735[⋆]	-0.284 [0.018]	-0.314[⋆]	-0.344 [⋆]
Residence	Urban				
	Rural	-0.087[0.432]	-0.028[0.480]	-0.016 [0.721]	-0.001[0.922]
Oral contraceptive	NotUsed				
	used	-0.207[0.064]	-0.081[0.039]	-0.082[0.066]	-0.075[0.110]
Alcohol use	No				
	Yes	-0.324[0.004]	-0.139[0.001]	-0.149 [0.001]	-0.138[0.003]
Family history of breast cancer	No				
	With	0.239[0.032]	0.084[0.031]	0.091[0.041]	0.090[0.041]
Bfing	No				
	Yes	0.224[0.045]	0.094[0.017]	0.119[0.007]	0.118[0.011]
Recurrence	No				
	Yes	0.277[0.013]	0.102[0.009]	0.116[0.009]	0.116[0.013]
Tumor size	≤ 2 cm				
	2-5 cm	-0.121[0.047]	-0.106[0.044]	-0.121[0.039]	-0.121[0.047]
	> 5cm	-0.143[0.016]	-0.150[0.003]	-0.159[0.005]	-0.143[0.016]
Histologic grade	I				
	II	-0.280[0.063]	0.067[0.134]	0.084[0.111]	0.118[0.031]
	III	-0.346[0.017]	0.128[0.012]	0.164[0.004]	0.180[0.003]
Treatments taken	Chem				
	Rad	0.311[0.109]	0.128[0.059]	0.162[0.029]	0.177[0.023]
	Sur	0.640[0.030]	0.243[0.019]	0.273[0.011]	0.243[0.023]
	Hor	-0.019[0.925]	-0.025[0.722]	-0.018[0.820]	-0.022[0.794]
	Comb >2	-0.063[0.629]	-0.003[0.939]	0.022[0.667]	0.009[0.866]
Stages of breast cancer	I				
	II	-1.72[⋆]	-0.621[⋆]	-0.588[⋆]	-0.560[⋆]
	III	-2.25[⋆]	-0.807[]	-0.775[⋆]	-0.746[⋆]
	IV	-2.68[⋆]	-0.983[]	-0.978[⋆]	-0.946[⋆]
				⋆ is <0.0001	

Citation: **M.K. Sharma, Endalkachew Abebe.** *Determinants of Survival Time of Women with Breast Cancer. Archives of Oncology and Cancer Therapy. 2019; 2(2): 26-39.*