Survival Analysis of Breast Cancer Patients: A Case Study at Tikur Anbessa Specialized Hospital, Addis Ababa, Ethiopia

Asrat Demeke Wondimu

Department of Statistics, Dire Dawa University, Ethiopia

Abstract

Background: Breast cancer is one of the most common cancer diseases in women population worldwide. The aim of this study was to identify and examine the factors that affect the survival time of breast cancer patients.

Methods: This study was based on a cohort of 408 patients who had started their breast cancer treatments follow-up in 2011 at Tikur Anbessa Specialized Hospital. To model the survival time and examine the association between the survival times and predictors, the proportional hazard model with Weibull accelerated failure time was applied.

Results: Among the patients included in the study, the event (death) occurred for 14% of them while the remaining 86% were censored. The results of the analyses revealed that the survival time of breast cancer patients was significantly related with their age, family history of the disease, type of treatment taken, tumor size and recurrence of the disease.

Conclusion: The hazard of elderly patients, those with a family history of breast cancer, those with tumor size greater than 5 cm and those who experience local recurrence of the disease was found to be significantly higher. Medical experts need to specifically target such patients at any time of follow-up so as to minimize the risk of death due to breast cancer.

Keywords: *Survival analysis, Breast cancer, Cox proportional hazard model, Accelerated failure time*

1. Introduction

Chronic diseases like cancer and cardiovascular disease are major causes of morbidity and mortality. In particular, cancer is one of the most common and deadly diseases worldwide. Worldwide, the burden of disease impinges on the lives of tens of millions annually. According to GLOBOCAN 2008 estimates,

there were 12.7 million new cancer cases in 2008 worldwide, of which 5.6 million occurred in economically developed countries and 7.1 million in economically developing countries. The corresponding estimates for total cancer deaths in 2008 were 7.6 million, 2.8 million in economically developed countries and 4.8 million in economically developing countries (Ferlay et al., 2010).

By 2020, the total number of new cancer cases in the developing and developed world is predicted to increase by 73% and 29%, respectively, largely as a result of ageing, urbanization and change in dietary habits (Mathers, 1999). Despite this growing burden, cancer continues to receive low public health priority in Africa, largely because of limited resources and other pressing public health problems, including communicable diseases such as AIDS/HIV infection, malaria, and tuberculosis. It may also be in part due to lack of awareness about the magnitude of the current and future cancer burden among policy-makers, the general public, and international private or public health agencies (Parkin and Sanghvi, 1991).

The incidence of breast cancer varies between countries; the highest rates occur in the United States and Canada and, the lowest rate is found in Asia and African populations. Estimates suggest that more than one million new breast cancer cases occur worldwide annually, with nearly 580,000 cases occurring in developed countries. Thus, breast cancer now ranks first among cancers affecting women throughout the world and its marked impact is not restricted to Western industrialized societies (Jemal et al. 2011).

Breast cancer is the primary cause of cancer death among women globally. Breast cancer survival rates vary greatly worldwide, ranging from 80% or over in North America, Sweden and Japan to around 60% in middle-income countries and below 40% in low-income countries (Coleman et al., 2008). Early detection linked to appropriate treatment is currently the most effective strategy to reduce breast cancer mortality. The low survival rates in less developed countries can be explained mainly by the lack of early detection programs, resulting in a high proportion of women presenting with late-stage disease, as well as by the lack of adequate diagnosis and treatment facilities (Anderson et al., 2008).

In Western Europe, breast cancer incidence has reached more than 90 new cases per 100,000 women annually, compared with 30 per 100,000 in eastern Africa. In contrast, breast cancer mortality rates in these two regions are almost identical, at about 15 per 100,000, which clearly points to a later diagnosis and much poorer survival in eastern Africa (Ferlay, 2015).

In Ethiopia, no population-based prevalence study exists, but hospital records show that there are more than 200,000 cancer cases per year where cervical and breast cancers are the top two cancer types having a lion's share for the high maternal deaths in the country. Breast cancer is the second most often occurring cancer next to cervical cancer among women in Ethiopia. It is estimated that around 10,000 Ethiopian women and men have breast cancer with thousands of more cases unreported as women living in rural areas often seek treatment from traditional healers before seeking help from the government health system (Ethiopian Cancer Association, 2014).

A study utilizing the Kaplan-Meier method and Cox regression model involving a total of 133,057 female patients diagnosed with breast cancer in the United States indicated that younger breast cancer patients exhibit more aggressive disease than older patients. Middle-aged patients exhibit better overall survival (OS) and breast cancer-specific survival (BCSS) than young and elderly patients but exhibit BCSS rates similar to those of young patients after adjustments for confounders. Stratified analysis demonstrated that middle-aged patients exhibited better survival than young patients, with the exception of patients with stage III disease. An age of 60 years or more was a significant independent predictor of a poor prognosis (Chen et al., 2016). A number of studies also reported that elderly women experience poorer outcomes than younger patients (e.g., Schonberg et al., 2010).

Breast cancer risk is higher among women whose close blood relatives have this disease. Studies show that, in addition to well-established reproductive and lifestyle risk factors such as early age at menarche and HRT intake, there is a strong risk in relation to family history of breast cancer, with a twofold increase in risk of developing the disease for women with breast cancer in their first-degree family, and a larger increase in risk among women with a first-degree relative diagnosed before age 50 compared with after age 50 years (Barnard et al., 2015; Hemminki et al., 2002). According to a study on the early detection and diagnosis of breast cancer by Hider and Nicholas (1999), family history of breast cancer is a significant risk factor, being greater for those with first rather than second degree relatives.

Soerjomataram et al. (2008) showed that tumor size is one of the strongest prognostic indicators, even after 20 years of follow-up. A study by Chia et al. (2004) also found a lower 10-year overall survival rate for node-negative patients with a tumor of 2-5 cm compared to those with a tumor smaller than 1 cm (66% vs. 79%). Further, the study by Seedhom and Kamal (2011) revealed that tumor size was significantly associated with breast cancer survival: patients with tumor size of 5 cm and above had a higher risk of death than those with tumor size of 5 cm and less.

Other prognostic factors for the survival of breast cancer patients include the stage of breast cancer (Keyfitz and Littman, 1979), type of treatment (Rezaianzadeh et al., 2009) and local recurrence after treatment (Dunst et al., 2001).

There are only few studies that analyzed the prognostic factors for the survival of breast cancer patients using accelerated failure time (AFT) models in Ethiopia. Thus, the aim of this study was to identify and examine the factors that affect the survival time of breast cancer patients who got treatment at Tikur Anbessa hospital, Addis Ababa, Ethiopia, using AFT models.

2. Materials and Methods

2.1 Description of the study area and data

Tikur Anbessa Specialized Hospital is the largest general public hospital located in Addis Ababa. The Federal Ministry of Health estimates that there could be more than 150,000 cancer cases in Ethiopia each year, but available data is limited. As the nation's sole cancer referral center, the hospital is treating only about one percent of these patients. The population of this study was all breast cancer patients who had been registered at Tikur Anbessa Specialized Hospital in the year 2011 and under follow-up.

The computer-aided registry at the hospital contains information on patients' card number, name, sex and region. The card number of patients was used to find patients' follow up cards form the card room. The cards are prepared by the Federal Ministry of Health to be uniformly used by clinicians to early identify and document clinical and laboratory variables. Thus, the data on the variables considered in the study were collected from patients' follow up records.

Ethical permission was obtained from the Oncology Department of the hospital. Trained enumerators and the principal investigator collected the data from patient records. Secondary data from 416 patients available in the records have been considered in this study. However, eight of them had incomplete information and excluded from the analysis.

2.2 Variables of the study

The response variable for this study is time to event (death) of patients with breast cancer (measured in months). This is to say that the response variable is a censored survival time represented by variable time and event/death. Survival time refers to the follow-up time from a defined starting point to the occurrence of a given event. In case the event never happens by the time the study ends, the survival times are censored.

The predictor variables (factors) which are assumed to influence the survival of breast cancer patients were gender (Female, Male), Age, Region of the patients (Amhara, Tigray, Oromia, SNNP, Addis Ababa and Others), Family history of breast cancer, Treatment taken (surgery, chemotherapy, radiotherapy, a combination of treatments), Tumor size, Stage of breast cancer and Recurrence.

2.3 Methods of data analysis

To address the research questions of the study, different survival data analysis techniques were employed. A brief description of the survival data analysis methods considered in this study are presented below.

2.3.1 The Kaplan-Meier estimator

A survival function measures the probability of non-event after a certain time and is defined as:

$$S(t) = P(T > t) \quad \ldots\ldots (1)$$

where t denotes time and T is a random variable denoting the time of an event (failure). The survival function is always between 0 and 1. It is non-increasing and approaches to zero as time goes to infinitely.

Let $F(t)$ be the cumulative distribution function of the random variable T. $S(t)$ and $F(t)$ can be related through:

$$S(t) = P(T > t) = 1 - P(T \le t) = 1 - F(t) \quad \ldots\ldots (2)$$

The Kaplan-Meier method, also called product limit estimator, is based on individual survival times and assumes that censoring is independent of survival time (that is, the reason an observation is censored is unrelated to the cause of failure). The Kaplan-Meier method can be used to estimate the survival curve from the observed survival time without the assumption of an underlying probability distribution. The Kaplan-Meier estimator of survival at time t is defined as:

$$\hat{S}(t) = \prod_{t_i \le t} \left(1 - \frac{d_i}{n_i}\right) \quad \ldots\ldots (3)$$

where d_i is the number of patients who died at time t_i and n_i is the number of patients at risk before t_i.

2.3.2 The hazard function

The hazard function is defined as the event rate at time t conditional on survival until time t or later:

$$h(t) = \lim_{\Delta t \to 0} \frac{P(t < T \le t + \Delta t \mid T > t)}{\Delta t} = \frac{f(t)}{S(t)} \quad \ldots\ldots (4)$$

2.3.3 Modeling survival data

Non-parametric methods (such as Kaplan-Meier, Log-rank test) can be useful in the analysis of a single sample of survival data, or more groups of survival times. However, in most medical studies that give rise to survival data, supplementary information will also be recorded on each individual. The non-parametric methods discussed so far are not suitable for such data set. In order to explore the relationship between the survival experience of individuals and explanatory variables, an approach based on statistical modeling can be used (Collett, 2003).

Through a modeling approach to the analysis of survival data, we can explore how the survival experience of a group of individuals depends on the values of one or more explanatory variables whose values have been recorded for each individual at the time origin. A variety of models and methods have been developed for doing this sort of survival analysis using either parametric (AFT model and PH model) or semi-parametric approaches.

2.3.4 The Cox proportional hazards model

The basic model for survival data to be considered in this study is the proportional hazard model. This model was proposed by Cox (1972) and has also come to be known as the Cox regression model. It is a semi-parametric model that parametrically specifies the functional relationship between the lifetime of individuals and their characteristics but leaves the actual distribution of lifetimes arbitrary.

David Cox's (1972) paper took a different approach to standard parametric survival analysis and extended the methods of the nonparametric Kaplan-Meier estimates to regression type arguments for life table analyses. Cox advanced to prediction of survival time in individual subjects by only utilizing variables covering with survival and ignoring the baseline hazard of individuals. This was done by making no assumptions about the baseline hazard of individuals and only assuming that the hazard functions of different individuals remained proportional and constant over time.

Cox proportional hazard model is usually written in terms of the hazard model formula. This model gives an expression for the hazard at time t for an individual with a given specification of a set of explanatory variables (denoted by X) and is generally given by:

$$h(t, X, \beta) = h_0(t) \exp(X'\beta) \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots (5)$$

where $h_0(t)$ is the baseline hazard function that characterizes how the hazard function changes as a function of survival time, X is a vector of values of the explanatory variables, and $\beta = (\beta_0, \beta_1, \ldots, \beta_k)'$ is a vector of unknown regression parameters that are assumed to be the same for all individuals which measures the influence of the covariates on the survival experience.

The Cox proportional hazards regression model is formulated as the hazard function which computes a coefficient for each of the predictor covariates that indicates the direction and degree of effect for the hazard of breast cancer. In a proportional hazards regression model, the unique effect of a unit increase in a covariate is multiplicative with respect to the hazard rate.

2.3.5 Accelerated failure time model

If the proportionality assumption is not valid, the Cox proportional hazard models cannot be used in modeling, rather some parametric approaches are appropriate. In a parametric model, the distribution of outcomes (time to event) is specified in terms of a finite number of unknown parameters. One of the famous parametric models is accelerated failure time (AFT) model in which the time to event is assumed to be a function of explanatory variables. The AFT model assumes that the relationship between the logarithm of T (log survival time) and X is linear. It can be written as an ordinary regression model for log survival time of the form:

$$Y = \log(T) = X'\beta + \sigma W \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots (6)$$

where $\sigma = 1/\sqrt{\tau}$ is scale parameter and W is the error term which has a suitable distribution, that is, $W \sim F_W(\cdot|\sigma)$ such that $F_W(\cdot|\sigma)$ is a known CDF associated with the density $f_W(\cdot|\sigma)$. The survival and hazard functions of W are $S_W(\cdot|\sigma) = 1 - F_W(\cdot|\sigma)$ and $h_W = f_W/S_W$, respectively.

The commonly used distributions for W are extreme value, normal and logistic and these lead to Weibull, log-normal and log-logistic models for T, respectively.

a) If W is extreme value, then T is Weibull, $T \sim \text{Weib}(\sqrt{\tau}, e^{-X'\beta\sqrt{\tau}})$, which is an exponential distribution whenever $\tau = 1$. It can be shown that the coefficients from the AFT and PH formats are linked in the following manner:

$$\hat{\beta}_{PH} = \frac{-\hat{\beta}_{AFT}}{\hat{\sigma}}$$

where $\hat{\beta}_{PH}$ and $\hat{\beta}_{AFT}$ are the estimated coefficient vectors for proportional hazard and AFT models, respectively.

b) If W is standard normal, then T is log normal, $T \sim LN(X'\beta, 1/\tau)$.

$$f(t \mid X, \beta, \tau) = \frac{1}{\sqrt{2\pi}} \frac{\sqrt{\tau}}{t} \exp\left\{-\frac{\tau}{2}[\log(t) - X'\beta]^2\right\}$$

c) If W has a logistic distribution, then T has log-logistic distribution, $T \sim LL(X'\beta, \sqrt{\tau})$.

Table 1 summarizes common error distributions for AFT models.

Table 1: Error distributions for Accelerated failure time models

Baseline distribution	$f_W(w)$	$F_W(w)$	Distribution of T
Extreme Value	$e^w \exp(-e^w)$	$1-\exp(-e^w)$	Weibull
Normal	$\frac{1}{\sqrt{2\pi}} e^{-w^2/2}$	$\Phi(w)$	Lognormal
Logistic	$e^w/(1+e^w)^2$	$e^w/(1+e^w)$	Log-logistic

3. Results and Discussion

The study aimed to identify the determinant risk factors for the survival of patients with breast cancer at Tikur Anbessa Specialized Hospital who started their treatment in the year 2011. In this study, a sample of 408 breast cancer patients of which 95.6% were female and 4.4% were male was considered. Among these patients, 86% and 14% were censored and died, respectively. When we see the region of the breast cancer patients, 46.8% were from Addis Ababa, 23.5% from Oromia, 11.0% from Amhara, 10.3% from SNNP and 4.9% from other regions. The death proportion was higher in Tigray region (28.6%).

3.1 Cox proportional hazard model and assessment of proportionality

We begin with a multivariable survival analysis of breast cancer patients using Cox proportional hazard model. It is well known that assessing whether the assumption of proportional hazards holds or not is a central theme in survival analysis. Multiple options for testing and accounting for non-proportionality are available. One of the methods for assessing departure from non-proportionality is by introducing time-by-covariate interactions to the Cox model, and test for their significance (Cox, 1972). In our case, the interactions of the covariates age, family history of breast cancer, treatment taken and recurrence with the

logarithm of time were found to be not significant – an indication that the proportional hazard assumption was satisfied for each of these covariates. However, the interaction terms for tumor size and treatment taken were significant. The implication is that the proportional hazard assumption is violated for these covariates.

Figure 1 presents the scatter plots of scaled Schoenfeld residuals from the conventional Cox model against time for the covariates treatment taken (left panel) and tumor size (right panel). We can observe that the distributions of the residuals are non-systematic from the reference line. This is a further indication of the violation of the proportional hazards assumption for these covariates. For instance, an increasing linear trend for the scaled Schoenfeld residuals of tumor size may be an indication of an increasing hazard ratio over time. Thus, we doubt the accuracy of the PH assumption and consider the AFT model for our data set.

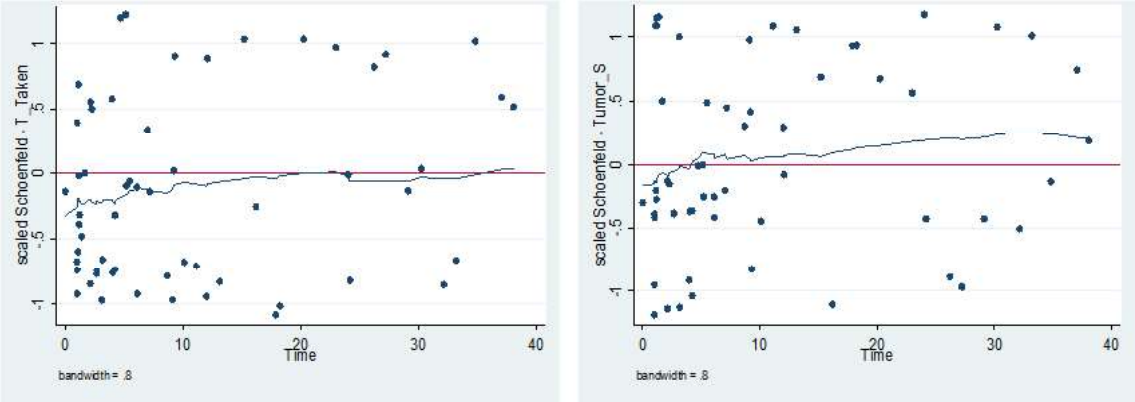

Figure 1: The plot of Scaled Schoenfeld residuals for treatment taken (left panel) and tumor size (right panel)

3.2 Accelerated failure time models

Due to the non-proportionality of some of the covariates, we fitted three AFT models for the survival of patients with breast cancer. In analyzing whether the assumed distribution best describes the data at hand, we can use the Cox-Snell residuals. If the model fits the data well, then these residuals follow the standard exponential distribution with mean one (Collett, 2003). In our case, the Cox-Snell residuals plots indicated that the Weibull regression model fits the data better than lognormal and log-logistic regression models. However, graphical methods for model choice may not assure the result. One of the common criteria used to select the best fitting model among candidate models is the Akaike information criterion (AIC) (Akaike, 1974). Table 2 presents the AIC for the three AFT models under consideration. We can

see that the Weibull regression model has the least AIC value. Thus, the data on patients with breast cancer better fits the Weibull AFT model.

Table 2: Comparison between parametric AFT regression models

Models	Weibull	Log-logistic	Lognormal
AIC	384.177	388.608	405.3798

3.3 Multivariable analysis of Weibull AFT model

Table 3 pertains to the fitted Weibull AFT model using forward stepwise (Conditional LR) variable selection method. It can be seen that the variables age, family history of breast cancer, type of treatment taken, tumor size and recurrence have significant influence on the survival probability of patients with breast cancer (p-value < 0.05).

Table 3: The results of the fitted Weibull AFT model (at Tikur Anbessa Specialized Hospital)

Variable	$\hat{\beta}$	SE	[95% CI]	P-value
Intercept	8.9791	1.0580	(6.905, 11.053)	0.000
Age	-0.0834	0.0145	(-0.112, -0.055)	0.000
Family history				
No (Ref.)				
Yes	-0.7959	0.3681	(-1.517, -0.074)	0.031
Treatment taken				
Chemotherapy (Ref.)				
Radiotherapy	-0.8523	0.3896	(-1.616, -0.089)	0.029
Surgery	-0.9847	0.4016	(-1.772, -0.198)	0.014
Combination of treatments	0.5726	0.3961	(-0.204, 1.349)	0.148
Tumor size				
T<=5 (Ref.)				
T>5	-1.7897	0.7956	(-3.349, -0.230)	0.024
T=T4	-1.1365	0.4112	(-1.942, -0.331)	0.006
Unknown	-0.2889	0.5283	(-1.324, 0.747)	0.584
Recurrence				
Yes (Ref.)				
No	0.9590	0.3712	(0.232, 1.687)	0.010
Log(scale)	0.0850	0.1103	(-0.276, 0.153)	0.443
Scale (σ)	1.0880	0.1200	(0.876,1.351)	

Ref. = Reference category

We can see from the results that a one unit increase in age decreased the survival time of breast cancer patients by a factor of $\exp(-0.0834) = 0.9199$. Using the relationship between the coefficients of AFT and PH models ($\hat{\beta}_{PH} = -\hat{\beta}_{AFT}/\hat{\sigma}$), for a one year increase in age, the hazard of a patient increased by

$\exp(0.0834/1.088)$ = 1.08 units. Our result is consistent with those of Chen et al. (2016) and Schonberg et al. (2010) who reported that elderly patients experience poorer outcomes than younger patients.

The hazard of a patient with family history of breast cancer was $\exp(0.7959/1.088)$ = 2.08 times that of patients who had no family history of breast cancer, keeping other factors constant. A number of studies also found that the risk of breast cancer is higher among women whose close blood relatives have this disease (e.g., Barnard et al., 2015; Hemminki et al., 2002; Hider and Nicholas, 1999).

The other significant prognostic factor for the survival of breast cancer patients was the type of treatment. Keeping the other covariates constant, the hazards of patients who took radiotherapy and surgery were $\exp(0.8523/1.088)$ = 2.19 and $\exp(0.9847/1.088)$ = 2.47 times that of patients who took chemotherapy, respectively. This finding is not in line with that of Rezaianzadeh et al. (2009) who reported that patients who underwent surgery as the first treatment option had a better prognosis than those who were treated firstly by chemotherapy possibly due to a larger tumor size.

Our result also revealed that tumor size was a significant predictor of the survival of breast cancer patients. The hazard of a patient with tumor size greater than 5 cm was 5.18 times that of a patient whose tumor size was less than or equal to 5 cm, keeping other covariates constant. This figure was 2.84 for a patient whose tumor was any size with direct extension to a chest wall (T4). The study by Seedhom and Kamal (2011) also found that tumor size was significantly associated with breast cancer survival. The study also found that patients with local recurrence after treatment of breast cancer had a higher risk of death than those without recurrence ($\exp(-0.9590/1.088)$ = 0.41).

4. Conclusion and Recommendations

This study conducted at Tikur Anbessa Specialized Hospital in Addis Ababa, Ethiopia, investigated the risk factors associated with breast cancer survival in women. Among the patients included in the study, the event (death) occurred for 14% of them while the remaining 86% were censored.

When fitting a multivariable Cox model by introducing time-by-covariate interactions, significant time-by-covariate interactions involved tumor size and treatment taken. This result is an indication that the hazard ratios associated with these risk factors were not constant over time, and hence, the proportional hazard model was not appropriate. Accordingly, AFT models were considered for analysis. Among the

candidate AFT models, the Weibull regression model was found to be a better fit since it had the least AIC value.

The fitted Weibull AFT model revealed that patient characteristics such as age, tumor size, type of treatment, family history of the disease and disease recurrence were significant risk factors associated with the survival time of women with breast cancer. Medical experts could make use of these factors when handling their breast cancer patients at any time of follow-up so as to minimize the risk of death due to breast cancer.

References

1. Akaike, H. (1974). A New Look at the Statistical Model Identification. *IEEE Transactions on Automatic Control*, 19:716–723.
2. Anderson BO., Yip CH., Smith RA., et al. (2008). Guideline Implementation for Breast Healthcare in Low-income and Middle-income Countries: Overview of the Breast Health Global Initiative Global Summit 2007. *Cancer*, 113(8):2221–43.
3. Barnard ME., Boeke CE. and Tamimi RM. (2015). Established Breast Cancer Risk Factors and Risk of Intrinsic Tumor Subtypes. *Biochim Biophys Acta*., 1856(1):73-85.
4. Chen HL., Zhou MQ., Tian W., Meng KX. and He HF. (2016). Effect of Age on Breast Cancer Patient Prognoses: A Population-Based Study Using the SEER 18 Database. *PLoS One*, 11(10).
5. Chia SK., Speers CH., Bryce CJ., et al. (2004). Ten-Year Outcomes in a Population-Based Cohort of Node-Negative, Lymphatic, and Vascular Invasion-Negative Early Breast Cancers without Adjuvant Systemic Therapies. *Journal of Clinical Oncology*, 22:1630-1637.
6. Coleman PM., Quaresma M., Berrino F., Lutz JM., et al. (2008). Cancer Survival in Five Continents: A Worldwide Population-Based Study (CONCORD). *Lancet Oncology*, 9:730-56.
7. Collett, D. (2003). Modelling Survival Data in Medical Research. Chapman & Hall, London.
8. Cox, D. (1972). Regression Models and Life-Tables. *Journal of the Royal Statistical Society, Series B*, 34:187-220.
9. Dunst J., Steil B., Furch S., Fach A., Lautenschläger C., Diestelhorst A., Lampe D., Kölbl H. and Richter C. (2001). Prognostic Significance of Local Recurrence in Breast Cancer After Postmastectomy Radiotherapy. *Strahlenther Onkol*., 177(10):504-10.
10. Ethiopian Cancer Association (ECA). (2014): Available from: http://ccethiopia.org/index.html.

11. Ferlay, J., Shin, H., Bray, F., Forman, D., Mathers, C. and Parkin, DM. (2010). Estimates of Worldwide Burden of Cancer in 2008: GLOBOCAN 2008. *International Journal of Cancer*, 127(12):2893-2917.
12. Ferlay J., Soerjomataram I., Dikshit R., Eser S., Mathers C., Rebelo M., Parkin DM., Forman D. and Bray F. (2015). Cancer Incidence and Mortality Worldwide: Sources, Methods and Major Patterns in GLOBOCAN 2012. *International Journal of Cancer*, 136(5):E359-86.
13. Hemminki K., Granström C. and Czene K. (2002). Attributable Risks for Familial Breast Cancer by Proband Status and Morphology: A Nationwide Epidemiologic Study from Sweden. *International Journal of Cancer*, 100(2):214-9.
14. Hider, P. and Nicholas, B. (1999). The Early Detection and Diagnosis of Breast Cancer: a Literature Review. New Zealand Health Technology Assessment (NZHTA) report. Volume 2, no. 2. Christchurch, N.Z.
15. Jemal A., Bray F., Center MM., Ferlay J., Ward E. and Forman D. (2011). Global Cancer Statistics. *CA: A Cancer Journal for Clinicians*, 61(2):69-90.
16. Keyfitz N. and Littman G. (1979). Mortality in a Heterogeneous Population. *Population Studies*, 33:333–42.
17. Mathers C., Bernard C., Iburg K., et al. (1999). Global Cancer Statistics. *CA: A Cancer Journal for Clinicians*, 49:33–64.
18. Parkin DM. and Sanghvi LD. (1991). Cancer Registration in Developing Countries. In: Jensen OM, Parkin DM, MacLennan R, Muir CS, Skeet RG, eds. Cancer Registration: Principles and Methods. Lyon, IARC Scientific Publications, 185–198.
19. Rezaianzadeh, A., Peacock, J., Reidpath, D., et al. (2009). Survival Analysis of 1148 Women Diagnosed with Breast Cancer in Southern Iran. *BMC Cancer*, 9:168.
20. Schonberg MA., Marcantonio ER., Li D., Silliman RA., Ngo L. and McCarthy EP. (2010). Breast Cancer among the Oldest Old: Tumor Characteristics, Treatment Choices, and Survival. *Journal of clinical oncology*, 28(12):2038–45.
21. Seedhom AE. and Kamal NN. (2011). Factors Affecting Survival of Women Diagnosed with Breast Cancer in El-Minia Governorate, Egypt. *International Journal of Preventive Medicine*, 2(3):131-8.
22. Soerjomataram I., Louwman MW., Ribot JG., Roukema JA. and Coebergh JW. (2008). An Overview of Prognostic Factors for Long-term Survivors of Breast Cancer. *Breast Cancer Research and Treatment*, 107(3):309-30.