

Parametric survival model to identify the predictors of breast cancer mortality: An accelerated failure time approach

Zeinab Irajil¹, Tohid Jafari Koshki¹, Roya Dolatkah², Mohammad Asghari Jafarabadi^{1,3}

¹Department of Statistics and Epidemiology, Faculty of Health, Tabriz University of Medical Sciences, Tabriz, Iran, ²Hematology and Oncology Research Center, Tabriz University of Medical Sciences, Tabriz, Iran, ³Road Traffic Injury Research Center, Tabriz University of Medical Sciences, Tabriz, Iran

Background: Breast cancer (BC) was the fifth cause of mortality worldwide in 2015 and second cause of mortality in Iran in 2012. This study aimed to explore factors associated with survival of patients with BC using parametric survival models. **Materials and Methods:** Data of 1154 patients that diagnosed with BC recorded in the East Azerbaijan population-based cancer registry database between March 2007 and March 2016. The parametric survival model with an accelerated failure time (AFT) approach was used to assess the association between sex, age, grade, and morphology with time to death. **Results:** A total of 217 (18.8%) individuals experienced death due to BC by the end of the study. Among the fitted parametric survival models including exponential, Weibull, log logistic, and log-normal models, the log-normal model was the best model with the Akaike information criterion = 1441.47 and Bayesian information criterion = 1486.93 where patients with higher ages (time ratio [TR] = 0.693; 95% confidence interval [CI] = [0.531, 0.904]) and higher grades (TR = 0.350; 95% CI = [0.201, 0.608]) had significantly lower survival while the lobular carcinoma type of morphology (TR = 1.975; 95% CI = [1.049, 3.720]) had significantly higher survival. **Conclusion:** Log-normal model showed to be an optimal tool to model the survival of patients with BC in the current study. Age, grade, and morphology showed significant association with time to death in patients with BC using AFT model. This finding could be recommended for planning and health policymaking in patients with BC. However, the impact of the models used for analysis on the significance and magnitude of estimated effects should be acknowledged.

Key words: Accelerated failure time, breast cancer, parametric model, survival

How to cite this article: Irajil Z, Jafari Koshki T, Dolatkah R, Asghari Jafarabadi M. Parametric survival model to identify the predictors of breast cancer mortality: An accelerated failure time approach. *J Res Med Sci* 2020;25:38.

INTRODUCTION

Breast cancer (BC) is one of the prevalent cancers that is created by growth and spread of malignant cells from the breast tissue.^[1] It is a common cancer among women and the most common reason of cancer-related mortality.^[2] BC has the highest prevalence in the USA and Western Europe and the lowest prevalence in east Asia.^[2] Despite significant reductions in BC prevalence and mortality in many countries, its global prevalence is increasing, and its age-standardized incidence rate increase 12% annually.^[2-5]

The prevalence and mortality of BC are increasing in Iran and East Azerbaijan province in the northwest of Iran as well. Even though the BC incidence rates in East Azerbaijan are lower than the USA and Western Europe, the increase in BC incidence rates urges further study.^[6,7]

BC has complex etiology with different risk factors such as heredity, hormonal factors, environmental factors, nutrition, number of pregnancy, age, tumor grade, and morphology.^[8-10] Assessing the risk factors and proper planning are important to reduce BC incidence and mortality.

This is an open access journal, and articles are distributed under the terms of the Creative Commons Attribution-NonCommercial-ShareAlike 4.0 License, which allows others to remix, tweak, and build upon the work non-commercially, as long as appropriate credit is given and the new creations are licensed under the identical terms.

For reprints contact: reprints@medknow.com

Access this article online	
Quick Response Code: 	Website: www.jmsjournal.net
	DOI: 10.4103/jrms.JRMS_743_19

Address for correspondence: Prof. Mohammad Asghari Jafarabadi, Department of Statistics and Epidemiology, Faculty of Health, Tabriz University of Medical Sciences, Golgasht St., Attar e Neshabouri St., Postal Code: 5166614711, Tabriz, Iran. E-mail: m.asghari862@gmail.com

Submitted: 03-Nov-2019; **Revised:** 03-Dec-2019; **Accepted:** 11-Jan-2020; **Published:** 13-Apr-2020

Prediction of patients' survival and identifying high-risk patients are important for health management.^[11] Parametric survival models, among the various predictive models, use data distribution information to smooth the fluctuation caused by sampling error and may provide better estimates than those obtained by Cox model. Some of the parametric models use accelerated failure time (AFT) approach that directly targets the patients' survival probabilities as an important clinical quantity. These models do not require proportional hazards (PH) assumption which may not be satisfied in most practical settings.^[12] Hence, using AFT models could provide better estimates and fit in some situations. More optimal and exact estimates could provide a better measure for decision-making by the users of the results of this study.

Based on an extensive search in the literature, there was no study, if any, to investigate the predictors of survival in patients with BC, so in this study, we used some parametric survival models with AFT approach to investigate the relationship between the survival of patients with BC and common prognostic factors.

MATERIALS AND METHODS

Study population

The data of patients diagnosed with primary BC were used, which have been collected from March 20, 2007, to March 19, 2016, by the population-based cancer registry of East Azerbaijan province of Iran. The BCs have been confirmed according to C50.0–9 codes of the International Classification of Disease-Oncology for the patients.^[13]

Data collection

In the East Azerbaijan population-based cancer registry (EA-PBCR), all clinic-pathologic data were registered regularly including age, sex, morphology (including ductal carcinoma [DC]; lobular carcinoma [LC]; and other types), and Grade (I–IV) using 11-digit personal identification number. Information about survival and outcome was obtained by contacting patients or their relatives or referring to the hospital information system during follow-up.

Statistical analysis

Data were expressed as frequency (percentages) for categorical and mean \pm standard deviation for numeric variables. Survival, the primary outcome of our study, was calculated as the time from diagnosis of BC to death from BC or being censored by the end of the study period.

The PH assumption was checked for variables, and since it was not satisfied for grade, the parametric models were applied. We fitted parametric models including exponential (survival function defined as $S(t) = \exp[-\lambda t]$),

Weibull ($S(t) = \exp[-\lambda t^\alpha]$), log logistic ($S(t) = \frac{1}{1 + \lambda t^\alpha}$), and log-normal ($S(t) = 1 - \Phi\left[\frac{\ln t - \mu}{\sigma}\right]$). In AFT models, time ratio (TR) (sometimes called acceleration factor) is usually reported to reflect the impact of each variable on the survival. If TR for a variable is α , it means that the lifespan of this group, on average, is stretched out α -times longer than the lifespan of the reference group. AFT models assumed that the hazard of an event is constant over time. The assumption was assessed by the graphs of hazards over time. Having a constant increase in hazard overtime for all predictors confirmed this assumption.

All the variables were entered simultaneously in the multivariable model to obtain adjusted effects of the factors on the survival of patients. Statistical significance was set at 0.05. The fit of models was compared using probability plots and Akaike information criterion (AIC) and Bayesian information criterion (BIC), smaller value of which indicate better fit.

In addition, a sensitivity analysis was conducted to assess the effect of low number on men's in the study sample, so that the above mentioned optimal model was performed just on the women's data.

All analyses were performed using STATA MP 15.0 (StataCorp. 2017. Stata Statistical Software: Release 15. College Station, TX: StataCorp LLC).

RESULTS

Participant

A total of 1132 patients were female (98.1%) and mean age of patients was 50.4 (SD 12.5) years. The median (25th, 75th percentiles) of survival time was 46.83 (32.57, 69.28) months. A total of 217 (18.8%) individuals experienced death due to BC by the end of the study. The 1-, 3-, and 5-year survival rate were 96.1% (95% confidence interval [CI]: 94.81–97.07), 88.2% (95% CI: 86.10–89.97), and 81.4% (95% CI: 78.55–83.85), respectively.

Comparing mortality rate

The mortality rate from BC was 18.6% in females and 27.3% in males (P for difference = 0.305). The 1-, 3-, and 5-year survival rate was 96.2% (95% CI: 94.91–97.17), 88.2% (95% CI: 86.13–90.03), and 81.4% (95% CI: 78.58–83.93), respectively, in females and 91.9% (95% CI: 68.30–97.65), 86.4% (95% CI: 63.44–95.39), and 79.2% (95% CI: 52.35–91.91), respectively, in males.

Total percentage of BC death was 16.5 within patients aged < 50 years and 21.2 within patients aged \geq 50 years. The 1-, 3-, and 5-year survival rate was 97.6%

(95% CI: 95.98–98.57), 90.1% (95% CI: 87.24–92.33), and 84.4% (95% CI: 80.62–87.44), respectively, for patients aged <50 years and 94.6% (95% CI: 92.37–96.15), 86.2% (95% CI: 83.03–88.88), and 78.2% (95% CI: 73.84–82.00), respectively, for patients aged ≥50 years.

Furthermore, the mortality rate from BC was 2.1 within patients in Grade I, 14.7 within patients in Grade II, 26.9 within patients in Grade III, and 28.2 within patients in Grade IV. Furthermore, this rate was 17.8 within patients with DC type of morphology, 14.1 within patients with LC type and 31.1 within patients with other types.

Assessing the proportional hazards assumption

The PH assumption was checked for the variables, and it was confirmed for sex ($P = 0.35$), age ($P = 0.31$) and morphology ($P = 0.12$ and 0.49 , respectively, for DC and other types of morphology), and was not confirmed for grade ($P = 0.36, 0.04$, and 0.02 , respectively, for Category II, III, and IV).

Comparing the parametric models

According to AIC and BIC values [Table 1] and probability plots for survival data [Figure 1], the survival model with log-normal distribution had the best fit on the data.

Table 1: Akaike information criterion and Bayesian information criterion values for parametric models

Distribution	AIC	BIC
Exponential	1460.79	1501.20
Weibull	1460.05	1505.52
Log logistic	1454.37	1499.83
Log-normal	1441.47	1486.93

AIC=Akaike information criterion; BIC=Bayesian information criterion

The cumulative hazards of patients with BC are plotted against survival time in Figure 2.

Figure 2a indicates that the cumulative hazard for males is higher than females, but the difference was not significant ($P = 0.505$). As shown in Figure 2b, the cumulative hazard for patients aged ≥50 years is higher than that one for patients aged <50 years, and the difference was significant ($P = 0.021$). The differences in cumulative hazard among the morphology type groups were significant ($P < 0.001$). Figure 2c shows that the cumulative hazard for patients in LC type of morphology is lower than DC type, and the patients in other types of morphology had the highest cumulative hazard. The difference in cumulative hazard among the tumor grade groups was significant ($P < 0.001$), where it was the lowest for Grade I and the highest for Grade III and Grade IV [Figure 2d] and the difference between Grade III and Grade IV was not significant ($P = 0.464$).

According to the results of parametric regressions, higher age (≥50), higher grades and DC, and other morphologies (compared to LC) had significantly lower survival in BC patients [Table 2].

The results of sensitivity analysis to consider the effect of low number on men’s in the study sample were presented in Table 3. The results did not differ to those of analyses by total sample, and approximately the same values of TRs have been observed for risk factors.

DISCUSSION

In this study, we investigated the underlying factors of survival in patients diagnosed with primary BC through

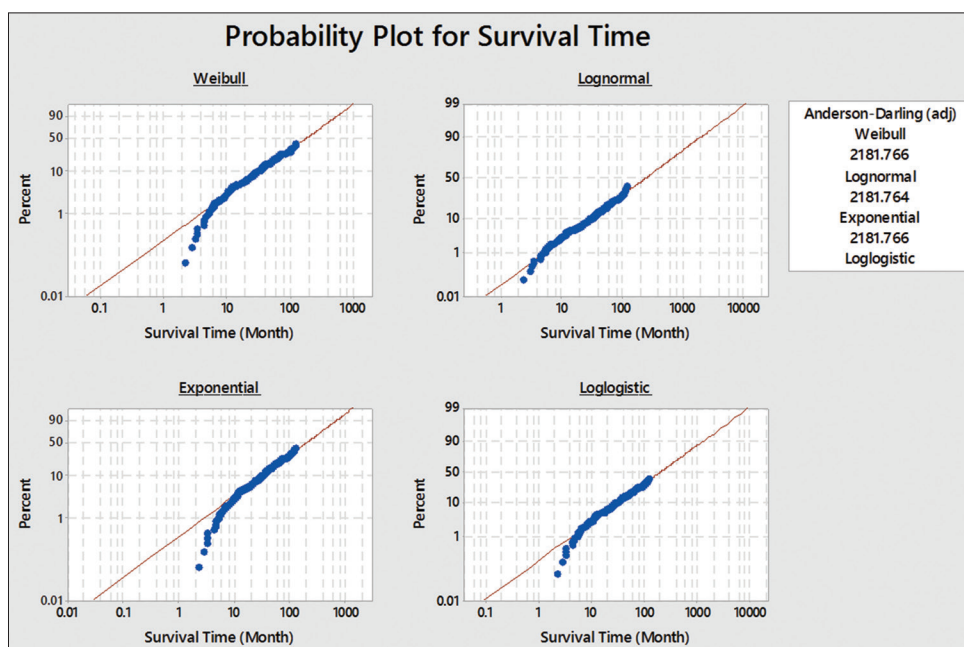


Figure 1: Probability plots for breast cancer data

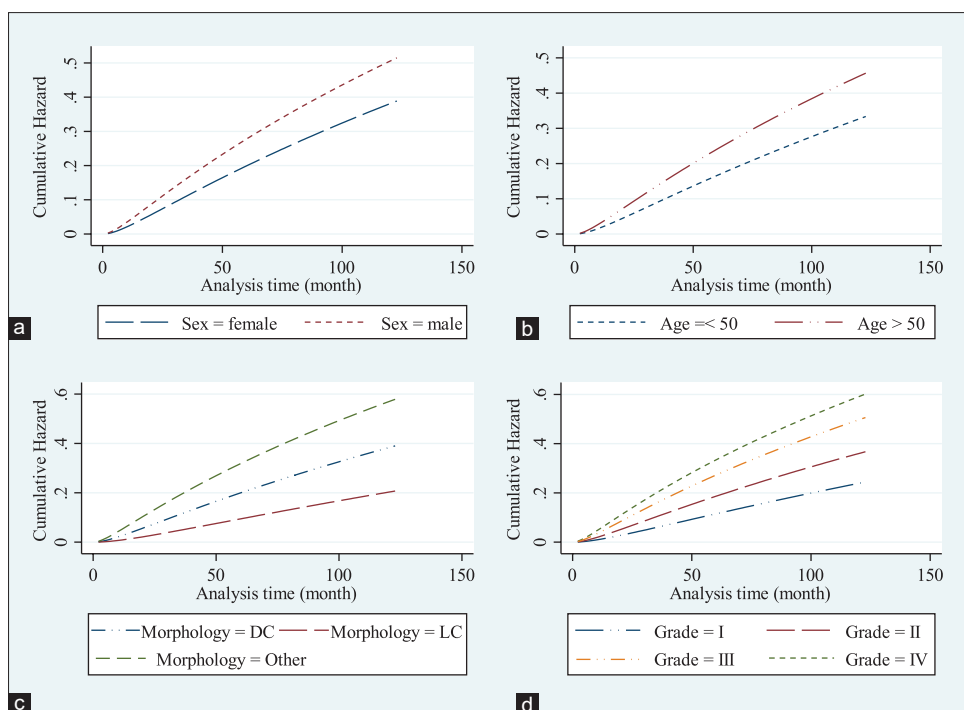


Figure 2: Cumulative hazard of patients with breast cancer by sex (a), no significant difference between males and females: $P=0.505$, by age (b), significant difference between higher and lower age: $P=0.021$, by morphology (c), significant difference among morphology types: $P<0.001$ and by grade (d), significant difference among different grades: $P<0.001$

Table 2: Parameter estimates and 95% confidence interval for log-normal accelerated failure time model

Variable	Exponential		Weibull		Log logistic		Log normal	
	Time ratio (95% CI)	P	Time ratio (95% CI)	P	Time ratio (95% CI)	P	Time ratio (95% CI)	P
Age								
<50	Reference		Reference		Reference		Reference	
≥50	0.72 (0.55-0.95)	0.019*	0.74 (0.58-0.95)	0.018*	0.72 (0.56-0.93)	0.012*	0.69 (0.53-0.90)	0.007*
Sex								
Female	Reference		Reference		Reference		Reference	
Male	0.81 (0.36-1.84)	0.617	0.83 (0.40-1.75)	0.635	0.80 (0.35-1.82)	0.590	0.71 (0.30-1.69)	0.439
Grade								
I	Reference		Reference		Reference		Reference	
II	0.65 (0.42-1.01)	0.057	0.67 (0.45-0.99)	0.047*	0.65 (0.44-0.97)	0.035*	0.64 (0.43-0.96)	0.031*
III	0.49 (0.31-0.76)	0.002*	0.53 (0.35-0.80)	0.003*	0.49 (0.32-0.74)	0.001*	0.44 (0.28-0.67)	<0.001*
IV	0.38 (0.22-0.66)	0.001*	0.42 (0.25-0.69)	0.001*	0.38 (0.23-0.66)	<0.001*	0.35 (0.20-0.61)	<0.001*
Morphology								
DC	Reference		Reference		Reference		Reference	
LC	1.72 (0.89-3.34)	0.106	1.63 (0.89-2.97)	0.114	1.76 (0.96-3.25)	0.068	1.97 (1.05-3.72)	0.035*
Other	0.65 (0.43-0.98)	0.039*	0.66 (0.45-0.96)	0.029*	0.65 (0.43-0.97)	0.038*	0.61 (0.40-0.95)	0.030*

* $P<0.05$. CI=Confidence interval; DC=Ductal carcinoma; LC=Lobular carcinoma

a parametric survival modeling approach. Interestingly, in our data, log-normal model had the best performance among all investigated parametric models. In line with our finding, log-normal model has outperformed other parametric models in modeling of progression-free survival in patients with advanced BC.^[14,15] While Baghestani *et al.* concluded that Weibull model is a better model in modeling the survival in patients with colorectal cancer.^[16]

Parametric survival models may offer optimal models due to using information of data distribution. Furthermore,

parametric survival models provide smooth graphs for survival and hazard functions and hence better estimation of such quantities specially in small sample studies or in the case of times with sparse data.^[12,17] Although our data were right-censored, parametric survival models can easily accommodate left- and interval-censored data as well.^[12] Furthermore, the PH assumption is not required to be assessed in some parametric model such as log-normal, log-logistic and generalized gamma do not depend on the PH assumption. In addition, in parametric models, the parameters can be estimated and completely

Table 3: Parameter estimates and 95% confidence interval for log-normal accelerated failure time model in women

Variable	Time ratio (95% CI)	P
Age		
<50	Reference	
≥50	0.67 (0.51-0.88)	0.004*
Grade		
I	Reference	
II	0.60 (0.40-0.91)	0.017*
III	0.40 (0.26-0.62)	<0.001*
IV	0.32 (0.18-0.56)	<0.001*
Morphology		
DC	Reference	
LC	2.32 (1.20-4.48)	0.012*
Other	0.65 (0.42-1.02)	0.060*

*P<0.05. CI=Confidence interval; DC=Ductal carcinoma; LC=Lobular carcinoma

specify the survival and hazard functions. This simplicity and completeness are the other main appeals of using a parametric approach.^[12]

Like our study, Foerster *et al.* found that sex had no significant relationship with survival.^[18]

Moreover, the age was inversely related to the survival of patients with BC. Previous studies have similarly revealed that the survival in patients with BC was correlated with their age.^[19] Correspondingly another study indicated that the risk of mortality had been increased by patients' age.^[9,20,21] Old patients may have lower survival than younger patients due to the existence of other possible persistent diseases, inappropriate health status, and diagnosis in a higher stage.^[20] On the other hand, some other studies did not indicate any significant relationship between patients' survival and their age.^[22-24]

The grade was another significant-related factor with survival so that an inverse relationship was observed between this factor and survival of patients. Similarly, previous studies have also discovered that patients' survival was inversely connected by the grade.^[9,19,22,25] It was concluded that, by increasing the tumor grade, the tumor becomes malignant, and therefore the patients' survival decreases.

Morphology was another significantly related factor of survival in patients with BC. The LC and DC types of morphology do not have any significant difference. The LC type had the highest survival compared to other types of morphology. Similar to our findings, Møller *et al.* found that LC type of morphology and Fallahzadeh *et al.* found that the DC and LC types of morphology had significantly higher survival in patients with BC.^[9,20]

Strengths and limitations

Obviously, our study was not without limitations. Some important clinical information such as tumor-node-metastasis stage, tumor size, estrogen receptor, and progesterone receptor were not available to be include in the analysis and improve the fit of models. As another limitation, in Iran, the Iranian National Cancer Registry and EA-PBCR have not yet developed. In the EA-PBCR, to avoid missing data, it was tried to collect data through a combined active and passive follow-up protocol and register all patients with newly diagnosed cancer from across the province. Because of the lack of contact information, the data for some of the initial cases were missed. This study was limited to recent 10 years because of the lag time in data reporting and registration.

CONCLUSIONS

Interestingly, using the parametric survival model, we achieved more flexibility to model the predictors of survival in patients with BC. Using log-normal parametric survival model led to remarkable advantageous AFT parameterization which could be interpreted based on survival (not hazard). The model showed that age, grade, and morphology were significant predictors of survival. These findings could be recommended for prevention, planning, health policymaking, early diagnosis of BC and early treatment and so increase survival in patients with BC.

Acknowledgments

Data of patients diagnosed with primary Breast cancer and registered in the East Azerbaijan population-based cancer registry were included in our analysis. As the ethics rules of East Azerbaijan population-based cancer registry, all patients' information, and records are confidential. The study protocol was approved by the Institutional Review Board of Tabriz University of Medical Sciences (IRB no.: IR.TBZMED.REC.1397.986).

Financial support and sponsorship

Research Deputy of Tabriz University of Medical Sciences, Tabriz, Iran.

Conflicts of interest

There are no conflicts of interest.

REFERENCES

1. Onsory K, Ranapoor S. Breast cancer and the effect of environmental factors involved. *New Cell Mol Biotechnol J* 2011;1:59-70.
2. Global Burden of Disease Cancer Collaboration, Fitzmaurice C, Allen C, Barber RM, Barregard L, Bhutta ZA, *et al.* 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;3:524-48.
3. Bray F, Ferlay J, Soerjomataram I, Siegel RL, Torre LA, Jemal A.

- Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA Cancer J Clin* 2018;68:394-424.
4. Fitzmaurice C, Fitzmaurice C, Allen C, Barber R, Barregard L, Bhutta Z, Brenner H, *et al.* Global, regional, and national cancer incidence, mortality, years of life lost, years lived with disability, and disability-adjusted life-years for 29 cancer groups, 1990 to 2016: A systematic analysis for the global burden of disease study. *JAMA Oncol* 2017;3:524-48.
 5. Somi MH, Dolatkhan R, Sepahi S, Belalzadeh M, Sharbafi J, Abdollahi L, *et al.* Cancer incidence in the East Azerbaijan province of Iran in 2015-2016: Results of a population-based cancer registry. *BMC Public Health* 2018;18:1266.
 6. Jafari-Koshki T, Schmid VJ, Mahaki B. Trends of breast cancer incidence in Iran during 2004-2008: A Bayesian space-time model. *Asian Pac J Cancer Prev* 2014;15:1557-61.
 7. Sharifian A, Pourhoseingholi MA, Emadedin M, Rostami Nejad M, Ashtari S, Hajizadeh N, *et al.* Burden of breast cancer in Iranian women is increasing. *Asian Pac J Cancer Prev* 2015;16:5049-52.
 8. Carlo JT, Grant MD, Knox SM, Jones RC, Hamilton CS, Livingston SA, *et al.* Survival analysis following sentinel lymph node biopsy: A validation trial demonstrating its accuracy in staging early breast cancer. *Proc (Bayl Univ Med Cent)* 2005;18:103-7.
 9. Møller H, Henson K, Lüchtenborg M, Broggio J, Charman J, Coupland VH, *et al.* Short-term breast cancer survival in relation to ethnicity, stage, grade and receptor status: National cohort study in England. *Br J Cancer* 2016;115:1408-15.
 10. Yang MT, Rong TH, Huang ZF, Zeng CG, Long H, Fu JH, *et al.* Clinical analysis of resectable breast cancer: A report of 6 263 cases. *Ai Zheng* 2005;24:327-31.
 11. Hothorn T, Lausen B, Benner A, Radespiel-Tröger M. Bagging survival trees. *Stat Med* 2004;23:77-91.
 12. Kleinbaum DG, Klein M. *Survival Analysis A Self-Learning Text*. New York, Springer-Verlag; 2001.
 13. Fritz A, Percy C, Jack A, Shanmugaratnam K, Sobin LH, Parkin DM, *et al.*, editors. *International Classification of Diseases for Oncology*. World Health Organization: WHO Library Cataloguing-in-Publication Data; 2013.
 14. Cope S, Ouwens MJ, Jansen JP, Schmid P. Progression-free survival with fulvestrant 500 mg and alternative endocrine therapies as second-line treatment for advanced breast cancer: A network meta-analysis with parametric survival models. *Value Health* 2013;16:403-17.
 15. Wang SJ, Kalpathy-Cramer J, Kim JS, Fuller CD, Thomas CR. Parametric survival models for predicting the benefit of adjuvant chemoradiotherapy in gallbladder cancer. *AMIA Annu Symp Proc* 2010;2010:847-51.
 16. Baghestani AR, Gohari MR, Orooji A, Pourhoseingholi MA, Zali MR. Evaluation of parametric models by the prediction error in colorectal cancer survival analysis. *Gastroenterol Hepatol Bed Bench* 2015;8:183-7.
 17. Klein JP, Moeschberger ML. *Survival Analysis: Techniques for Censored and Truncated Data*. New York, Springer-Verlag: Springer Science and Business Media; 2006.
 18. Foerster R, Foerster FG, Wulff V, Schubotz B, Baaske D, Wolfgarten M, *et al.* Matched-pair analysis of patients with female and male breast cancer: A comparative analysis. *BMC Cancer* 2011;11:335.
 19. Lotfnezhad Afshar H, Ahmadi M, Roudbari M, Sadoughi F. Prediction of breast cancer survival through knowledge discovery in databases. *Glob J Health Sci* 2015;7:392-8.
 20. Fallahzadeh H, Momayyezi M, Akhundzardeini R, Zarezardeini S. Five year survival of women with breast cancer in Yazd. *Asian Pac J Cancer Prev* 2014;15:6597-601.
 21. Fouladi N, Amani F, Harghi AS, Nayebyazdi N. Five year survival of women with breast cancer in Ardabil, North-West of Iran. *Asian Pac J Cancer Prev* 2011;12:1799-801.
 22. Haghghat S. Survival rate and its correlated factors in breast cancer patients referred to breast cancer research center. *Iran Q J Breast Dis* 2013;6:28-36.
 23. Karimi A, Delpisheh A, Sayehmiri K, Saboori H, Rahimi E. Predictive factors of survival time of breast cancer in Kurdistan province of Iran between 2006-2014: A cox regression approach. *Asian Pac J Cancer Prev* 2014;15:8483-8.
 24. Mohammadpour M, Yaseri M, Mahmoudi M, Entezar Mahdi R. Estimation of survival in women diagnosed with breast cancer with cure survival in West-Azerbaijan and East-Azerbaijan provinces. *J Sch Public Health Inst Public Health Res* 2018;16:63-74.
 25. Elsheikh SE, Green AR, Rakha EA, Powe DG, Ahmed RA, Collins HM, *et al.* Global histone modifications in breast cancer correlate with tumor phenotypes, prognostic factors, and patient outcome. *Cancer Res* 2009;69:3802-9.