Clinical Cancer Investigation Journal-ISSN: 2278-0513

Using Cure Models: A Five-year Survival Analysis of Breast Cancer Patients in Mazandaran Province

Abstract

The mortality rate has dropped as a result of recent considerable advancements in cancer therapies. Instead of the more common survival models, cure models are employed to analyze certain disorders. The current work uses Weibull models, which are adaptable cure models, to assess data from breast cancer patients.

The data set of the current descriptive study are from patients with breast cancer, who were referred to Imam Khomeini's hospital in Sari County during 2014-2017 and were followed in 2021. Individual characteristics of 221 patients were recorded overall using the census method. The data were analyzed using SPSS 23 and R 3.6.3 packages with the significance level set at 0.05.

The results showed that 99 (45%) of patient's deaths occurred after treatment. The rate of cure after one, three, and five years was found to be 0.97, 0.84, and 0.75 respectively. The results of the fitting model showed that Weibull's Non-Mixture Cure model (AIC=659) has a better fit compared to the other cure models. According to this model, factors affecting a patient's recovery include their age at diagnosis, smoking, metastasis, chemotherapy, and need for emergency hospitalization. Also, the cure rate was estimated to be 68%.

Weibull's Non-Mixture Cure model, modeling with Poisson distribution, and studying the effects of factors influencing the occurrence of death allow for a better analysis of the cancer trend and the provision of more precise data for researchers.

Keywords: *Survival analysis; Long-term survival; Breast Cancer; Cure models*

Bizhan Shabankhani[1], Jamshid Yazdani Charati[2], Gholamali Godazandeh[3], Milad Hadinia[4], Keihan Shabankhani[5], Saeed Kaviani Cherati[6]

1Assistant of Biostatic, Department of Biostatics, Faculty of Health, Mazandaran University of Medical Sciences, Sari, Iran (shabankhani@yahoo.com)

2Professor of Biostatic, Department of Biostatics, Faculty of Health, Mazandaran University of Medical Sciences, Sari, Iran (jamshid.charati@gmail.com)

3Associate Professor of Surgery, Department of Surgery, Medical School, Mazandaran University of Medical Sciences, Sari, Iran (ghgous@yahoo.com)

4MSc in Biostatistics, Student Research Committee, Faculty of Health, Mazandaran University of Medical Sciences, Sari, Iran (miladhadinia70@gmail.com)

5Department of Medicine, Medical School, Mazandaran University of Medical Sciences, Sari, Iran (kshabankhani@gmail.com)

6Department of Biostatistics, Faculty of Health, Mazandaran University of Medical Sciences, Sari, Iran (kaviani_ch93@yahoo.com)

Corresponding Author: Bizhan Shabankhani - Assistant of Biostatic, Department of Biostatics, Faculty of Health, Mazandaran University of Medical Sciences, Sari, Iran (Email: shabankhani@yahoo.com)

Introduction

The progress of human knowledge in recent decades has made us able to significantly decrease the rate of mortality and morbidity of some diseases and increase the average life span of human beings. However, the development of recently discovered diseases is a result of changes in lifestyle and their impact on the environment. Due to these factors, the prevalence of numerous diseases, including various cancers, has increased dramatically during the past few decades(1-3).

Breast cancer is one of the most common types of cancer among women. 2.3 million women in the world were affected by this disease and 685,000 women lost their lives only in 2020(4).

Breast cancer causes more disability-adjusted life years lost globally than any other type of cancer(5, 6). Breast cancer accounts for 76% of all cancer cases in Iranian women, with an estimated 40,000 patients overall and an average of 7,000 new cases every year(1).

Based on the results obtained, the rate of this disease is higher in developed countries than in developing countries, researchers consider this issue to be related to the youth of the population in these countries. But Researchers from the International Agency for Research on Cancer (IARC) and partner institutions estimate the impact this disease will have in 2040. Researchers predict that by 2040 the breast cancer burden will increase to more than 3 million new cases per year (an increase of 40%) and more than 1 million deaths per year (an increase of 50%)(7).

According to the report of the World Health Organization, the five-year survival rate after the diagnosis of this disease is 90% in high-income countries and 66% in other countries(4).

Along with the therapeutic advances in recent decades, data analysis methods have also made significant progress, which increases the accuracy of the results. As an example, the survival models are used in the analysis of data that patients experience a specific outcome (death) after a long period of illness such as cancer(8-10).

But today, thanks to improvements in treatment methods, not every patient dies as a result of their condition; instead, some may recover and become what is known as cured persons. For this reason, we use “Cure models” to analyze diseases that follow this structure(11-14).

By using the Cured models in the analysis of breast cancer patients, the survival of these patients can be estimated more accurately than the usual survival methods(11, 13, 15).

In a survival analysis, it is assumed that if the study's length is sufficient, every subject will experience the desired result. In some trials (like those involving breast cancer), a sizable fraction of participants may recover and never achieve the anticipated outcome (4). Even if we increase the duration of such studies, the probability of the outcome does not approach 1, for this reason, the usual survival estimates are biased and show misleading results. In such cases, it is better to use cured models(15-18).

Cured models are divided into two general categories, mixture cure models and non-mixture cure models.

Mixture cure models have three goals

1- Estimation of the proportion of cured people (they do not experience the expected outcome)
2- Estimating the survival function for individuals who experience the expected outcome
3-Factors affecting both exposed and cured groups

The non-mixture cure model assumes that the risk of an outcome occurring asymptotically tends to zero for all individuals. Unlike the mixture cure model, which assumes that all individuals are not exposed to the same outcome. As a result, we reach a proportion of people who have recovered(16).

This study aims to estimate the survival of breast cancer patients using mixture and non-mixture cure models in the presence of predictor variables.

Materials and Methods

Table 1: Information about the research variable

This study was a historical cohort study, which it’s statistical population is made up of women with breast cancer who continuously visited the Comprehensive Cancer Center of Imam Khomeini Hospital in Sari during the years 2014-2016. The Criteria for the study were the registration of a clinical diagnosis of cancer, and the required information was collected through file reading. Also, the patients were followed up with phone calls by two experienced interviewers for one month, in order to fill out the incomplete information of the patients regarding the outcome studied in the research. Finally, a census was conducted to record the data of 221 patients. These patients underwent a single modified radical mastectomy or breast preservation procedure, followed by adjuvant treatments like chemotherapy, radiation, and hormone therapy, with a minimum of five years of follow-up. The dependent variable was the time of the patient's death (the duration of the person's presence in the study from the time of diagnosis until death due to breast cancer or being censored) and the variables of age, marital status, Place of residence, economic status, education, family history, and surgery were considered as a covariate. The cure ratio of these patients was determined after treatment, and the people who were still alive at the end of the study period were considered cured people, and those whom we had no information about after a certain time were considered right censored. One of the limitations of this research was the lack of pathology information in the patient's medical records.

Due to the existence of cured people, the cure model (mixture, non-mixture) was used to analyze the data of this research, and the Weibull, exponential, log, and log-rank models were used for the distribution of the survival function of the patients. Akaike's information criterion (AIC) was used to select the most efficient fitted model. The lower the value, The better the proposed model. It should be noted that in this study, the Kaplan-Meier chart was used to identify the cured people. In this way, if the graph of the survival function continues as a horizontal and straight line before reaching zero, it will indicate the presence of healed people. For data analysis, SPSS version 23 and R version 3/6/3 software were used and a significance level of 5% was considered.

Results

In this study, 221 women with breast cancer were investigated. The average age of these patients at the time of initial diagnosis was 47.69 ± 11.03 years. 94 (42%) of the 221 patients died and 127 (58%) survived, which was considered as right censoring. Also, the descriptive statistics of other variables are given in the table below (Table 1).

Variable	Total Mean (n=221)	Censored mean (122)	Deceased mean (99)	Sig
Age at initial diagnosis	47.69±11.03	46.31±10.14	48.21±11.34	0.253
Place of habitation Village	76(34.39)	41(33.61)	35(35.35)	
City	145(65.61)	81(66.39)	64(64.65)	0.786
Smoking No	182(82.35)	105(86.07)	77(77.78)	
Yes	39(17.65)	17(13.39)	22(22.22)	0.108
Secondhand smoke No	90(40.72)	54(44.26)	36(36.36)	0.235
Yes	131(59.28)	68(55.74)	63(63.64)	
Occupation Housewife	218(98.64)	121(99.18)	97(97.98)	0.443
Employed	3(1.36)	1(0.82)	2(2.02)	
History of emergency No	144(65.16)	83(68.03)	61(61.62)	0.319
Hospitalization Yes	77(34.84)	39(31.97)	38(38.38)	
marital status Single	6(2.71)	3(2.46)	3(3.03)	1.000
Married	215(97.29)	119(97.54)	96(96.97)	
Involvement of lymph No	110(49.77)	74(60.66)	36(36.36)	<0.001
Nodes Yes	111(50.23)	48(39.34)	63(63.64)	
Mastectomy surgery No	68(30.77)	40(32.79)	28(28.28)	0.471
Yes	153(69.23)	82(67.21)	71(71.72)	
Family history of cancer No	146(66.06)	82(67.21)	64(64.65)	0.689
Yes	75(33.94)	40(32.79)	35(35.35)	
Chemotherapy No	49(19.91)	29(23.77)	15(15.15)	0.111
Yes	177(80.09)	93(76.23)	84(84.85)	
Underlying disease No	94(42.53)	54(44.26)	40(40.4)	0.564
Yes	127(57.47)	68(55.74)	59(59.6)	
Hormone Therapy No	22(9.95)	13(10.66)	9(9.09)	0.699
Yes	199(90.05)	109(89.34)	90(90.91)	
Radiotherapy No	55(24.89)	32(26.23)	23(23.23)	0.608
Yes	166(75.11)	90(73.77)	76(76.77)	
Economic status Poor	76(34.39)	37(30.33)	39(39.39)	0.158
Good	145(65.61)	85(69.67)	60(60.61)	

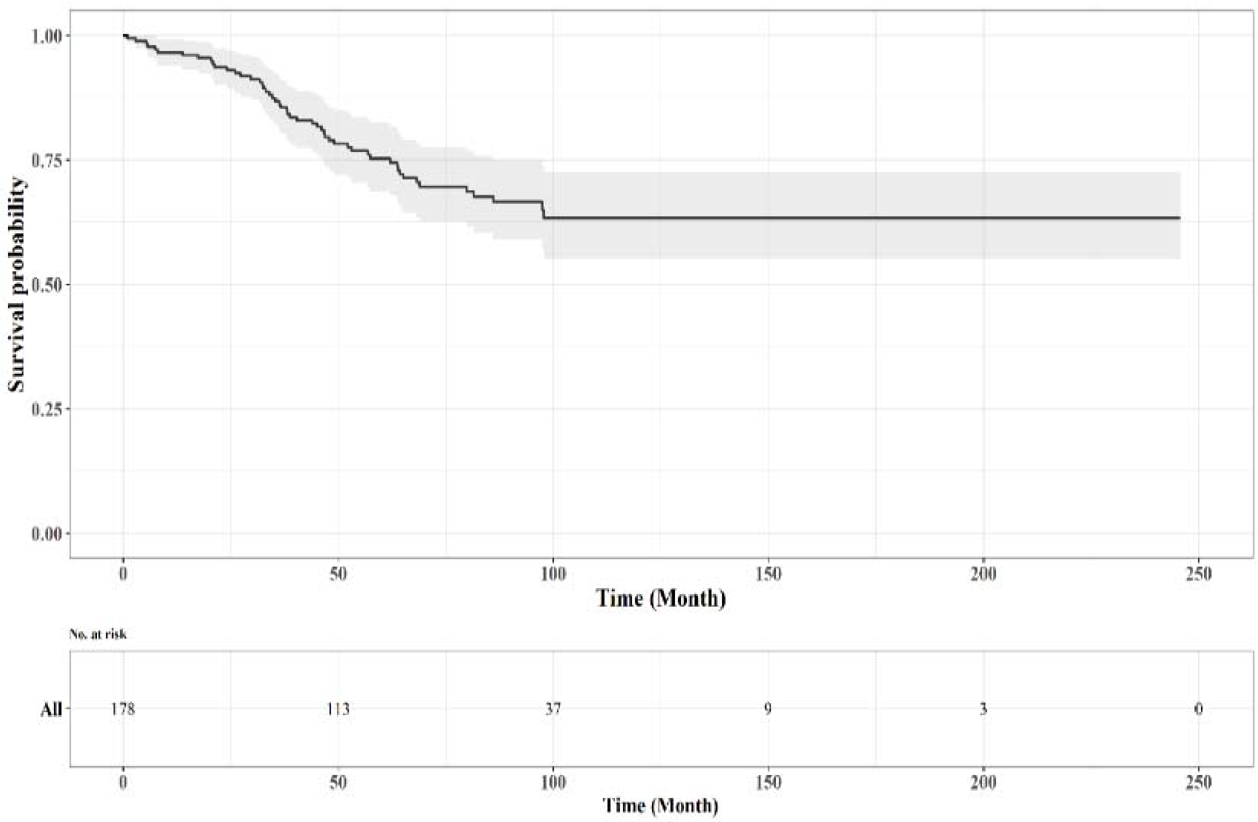
The survival rate in terms of the number of months was estimated as 97.74 (months). The point estimation for the survival rate of women with breast cancer is equal to 0.65 and the interval estimate of this rate is at the 95% C. I am equal to (0.57 - 0.74). Also, the one-, three-, and five-year survival rates were equal to 0.97, 0.84, and 0.75, respectively (Table 2).

Table2. Kaplan-Meier table to determine the survival of women with breast cancer

Time to event	Live at the start of the month	Number of deaths	Number of censors	survival time	Lower bound	Upper bound
1	178	0	0	1.00	1.00	1.00
10	166	6	6	0.97	0.94	0.99
20	160	2	4	0.95	0.92	0.99
30	148	7	5	0.91	0.87	0.96
40	130	12	6	0.84	0.78	0.89
50	113	8	9	0.78	0.72	0.85
60	97	4	12	0.75	0.69	0.82
70	83	7	7	0.70	0.63	0.78
80	68	1	14	0.69	0.62	0.77
90	55	2	11	0.67	0.59	0.75
97.74	40	1	14	0.65	0.57	0.74

The survival of more than five years is considered as cured and the Kaplan-Meier survival graph estimated from the data indicates the presence of cured people (flattening of the graph before reaching the zero point) in the study. Also, the community investigated in this study has been followed up for at least five years thus the follow-up period is sufficient (Figure 1).

Figure 1. Kaplan-Meier survival diagram estimated from the data



For the data of this research, mixture and non-mixture cure models with logit correlation function and Weibull, exponential, log-logistic, and log-normal distributions were used. The best model was selected through Akaike's criterion. Weibull distribution in both mixture and non-mixture models had the lowest Akaike value. Finally, Weibel's non-mixture cure model (AIC=659) has a better fit than other cure models (Table 3).

Table 3. Akaike's Information Criterion and estimated cure percentage for each of the distributions used in mixture and non-mixture cure models

Model	Survival function	Degree of freedom	Likelihood Log	Akaike information criterion	Estimating the percentage of curing (95% confidence interval)	Percentage of curing observed
Non-mixture	Weibel Exponantial Log Log-Normal	3 2 3 3	-0.327 -0.331 -0.328 -0.330	659 666 662 666	61.90(51.20-71.50) 37.30(8.22-79.80) 52.00(34.70-68.90) 19.95(0.39-94.11)	0.65(0.57-0.74) 0.65(0.74-0.74) 0.65(0.74-0.74) 0.65(0.74-0.74)
mixture	Weibel Exponantial Log Log-Normal	3 2 3 3	-0.327 -0.331 -0.328 -0.330	659 665 662 667	62.00(51.80-71.30) 46.66(23.03-71.88) 52.00(34.60-69.00) 28.66(1.61-90.82)	0.65(0.74-0.74) 0.65(0.74-0.74) 0.65(0.74-0.74) 0.65(0.74-0.74)

The estimated coefficients of predictor variables which are negative and their P-value is less than five percent, show that the cure rate is significantly lower for that predictor variable.

In Table 4, the variables of age, Smoking, involvement of lymph nodes, chemotherapy, and emergency hospitalization were significant.

Table 4. Fitting results of Weibull non-mixture cure models in the presence of predictor variables

Weibull non-mixture cure models					level	Variable
Sig	C.I	OR	B	Cured(%)		
- 0.046†	- (0.23-0.98)	- 0.48	- -0.74	0.68 0.50	<50* >50	Age
- 0.985	- (0.48-2. 13)	- 1.01	- -0.01	0.62 0.62	Village* Urban	place of habitation
- 0.762	- (0.64-2.72)	- 1.31	- -0.27	0.58 0.64	poor* good	Economic status
- 0.025†	- (0.16-0.88)	- 0.37	- -0.99	0.67 0.43	No* Yes	smoking
- 0.303	- (0.25-1.36)	- 0.33	- -0.47	0.68 0.57	No* Yes	secondhand smoker
- 0.985	- (0.48-2.07)	- 0.99	- -0.01	0.62 0.62	No* Yes	Family history of cancer
- 0.298	- (0.26-1.12)	- 0.54	- -0.62	0.70 0.56	No* Yes	History of underlying disease
- 0.001†	- (0.14-0.62)	- 0.30	- -1.21	0.74 0.46	No* Yes	History of emergency hospitalization
- 0.001†	- (0.06-0.30)	- 0.13	- -2.04	0.85 0.57	No* Yes	Involvement of lymph nodes
- 0.368	- (0.29-1.41)	- 0.64	- -0.45	0.59 0.57	No* Yes	Mastectomy
- 0.005†	- (0.05-0.58)	- 0.13	- -1.82	0.88 0.55	No* Yes	Chemotherapy
- 0.233	- (0.21-1.23)	- 0.50	- -0.69	0.74 0.74	No* Yes	Radiotherapy
- 0.223	- (0.12-1.65)	- 0.44	- -0.83	0.60 0.57	No* Yes	Hormone therapy

*Reference

†Sig(0.05)

Discussion

The common models used in the analysis of survival data are very wide but in most of these models, the main assumption is the occurrence of events that happen with the increase in the duration of the study. However, in some diseases with long-term observations, a number of subjects may not experience the main outcome and are considered right censored at the end of the study. We use cure models in such cases. One of the advantages of this model is providing a more realistic survival model in the absence of cured subjects in addition to estimating the proportion of cured people. It is worth noting that the results of this model are reliable only if the study time is long enough. One of the easiest ways to identify the cured subjects is to draw a Kaplan-Meier chart. If this graph reaches a flat surface before reaching zero, it will be a plausible reason for the existence of cured subjects(19).

In this study, the one, three, and five-year survival rates of patients are estimated to be 97%, 84,% and 75%, respectively. In Baghestani et al.'s study, one, five, and ten-year survival rates were estimated as 94%, 77,% and 56%, respectively(14). In another study conducted by Rezaianzadeh in 2009 in southern Iran, the three and five-year survival rates of patients were estimated at 76% and 58%, respectively.

The difference in five-year survival in the latter study with the other two studies can be attributed to the difference in the geography of the study site and the lifestyle of the people living there.

In this study, Weibel's non-mixture cure model with AIC=659 has the best fit compared to other cure models. The variables of age, smoking, involvement of lymph nodes, chemotherapy, and emergency hospitalization show a significant relationship with the reduction of healing in patients.

In Ghasemi et al.'s study, the size of the tumor and the severity of the tumor or stage? have a significant relationship with curing(20). That result was also confirmed in the study of Baghestani et al(21).

In a study conducted in 2018 by Nardin et al., the 5-year and 10-year survival rates of patients in Western countries were 90% and 80%, respectively. This high percentage is due to early diagnosis and innovative systemic treatments, which have a better trend in Western countries than in developing countries(22).

In a study by Haqua et al. in 2008 on 934 women who were diagnosed with breast cancer between 1988 and 1995, normal survival models were used but the results of this study show that after the 13th year of the diagnosis, the Kaplan-Meier diagram is completely horizontal. Therefore, the appropriate model is the cure model for conducting this study(23).

In another study conducted by Usman et al. in Nigeria on 312 women with breast cancer, the normal survival models were used. The results of this research showed that the Kaplan-Meier chart is completely horizontal after 35 months. Therefore, in this case, the cure models should be prioritized(24).

Conclusion

The results of this study and the mentioned studies and also the relationship of these factors with the longer survival of patients emphasize the importance of rapid breast diagnosis and timely treatment in these people. Better analysis of the cancer trend and the availability of more precise data for researchers are made possible by Weibull's Non-Mixture Cure model, modeling using Poisson distribution, and investigating the effects of factors influencing the occurrence of mortality.

Finally, in light of related research, it is important to focus on two fundamental issues. The first is patient survival, which is a quantity that depends on factors like diagnosis timing, type of treatment, lifestyle, etc. The statistical model that was used to analyze the data is the second point, particularly in research with a long enough study period. The cure models offer a better fit than the other survival models if the Kaplan-Meier curve flattens prior to the end of the research.

Statements and Declarations

The authors did not receive support from any organization for the submitted work.

Acknowledgments

We would like to thank Mazandaran University of Medical Sciences for giving us permission to conduct this study.

Conflict of interest

Dr. Bizhan Shabankhani is an assistant professor of Biostatistics at Mazandaran University of Medical Sciences and declares no relevant conflicts of interest. All authors declare that they have no other conflicts of interest.

Financial support

The authors did not receive financial support from any organization for this study.

Ethical statement

None.

….

References

1. Maetani S, Gamel JW. Parametric cure model versus proportional hazards model in survival analysis of breast cancer and other malignancies. Advances in Breast Cancer Research. 2013;2(4):119-25.

2. Babu GR, Samari G, Cohen SP, Mahapatra T, Wahbe RM, Mermash S, et al. Breast cancer screening among females in Iran and recommendations for improved practice: a review. Asian Pac J Cancer Prev. 2011;12(7):1647-55.

3. Kashyap D, Pal D, Sharma R, Garg VK, Goel N, Koundal D, et al. Global Increase in Breast Cancer Incidence: Risk Factors and Preventive Measures. BioMed Research International. 2022;2022:9605439.
4. MohammadSharifi A, Hajizadeh E, Baghestani A, Haji FathAli A. Mixture And Non-Mixture Cured Models In Survival Analysis of leukemia patients: A cohort study. Studies in Medical Sciences. 2021;31(10):802-12.
5. Ghorbani N, Yazdani Cherati J, Anvari K, Ghorbani N. Factors affecting recurrence in breast cancer using Cox model. Journal of Mazandaran University of Medical Sciences. 2015;25(131):32-9.
6. 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. Journal of School of Public Health and Institute of Public Health Research. 2018;16(1):63-74.
7. Arnold M, Morgan E, Rumgay H, Mafra A, Singh D, Laversanne M, et al. Current and future burden of breast cancer: global statistics for 2020 and 2040: WHO; 2022 [Available from: https://www.iarc.who.int/news-events/current-and-future-burden-of-breast-cancer-global-statistics-for-2020-and-2040/.
8. Adamu PI, Adamu MO, Okagbue HI, Opoola L, Bishop SA. Survival analysis of cancer patients in north eastern Nigeria from 2004–2017–a Kaplan-Meier method. Open Access Macedonian Journal of Medical Sciences. 2019;7(4):643-50.
9. Talebi A, Mohammadnejad A, Akbari A, Pourhoseingholi MA, Doosti H, Moghimi-Dehkordi B, et al. Survival analysis in gastric cancer: a multi-center study among Iranian patients. BMC surgery. 2020;20(1):1-8.
10. Tesfay B, Getinet T, Derso EA. Survival analysis of Time to Death of Breast Cancer Patients: in case of Ayder Comprehensive Specialized Hospital Tigray, Ethiopia. Cogent Medicine. 2021;8(1):1-11.
11. Amico M, Van Keilegom I. Cure models in survival analysis. Annual Review of Statistics and Its Application. 2018;5:311-42.
12. Burguin A, Diorio C, Durocher F. Breast Cancer Treatments: Updates and New Challenges. J Pers Med. 2021;11(8).
13. Zhai J, Wu Y, Ma F, Kaklamani V, Xu B. Advances in medical treatment of breast cancer in 2022. Cancer Innovation. 2023;2(1):1-17.
14. Baghestani AR, Shahmirzalou P, Sayad S, Akbari ME, Zayeri F. Comparison cure rate models by DIC criteria in Breast Cancer data. Asian Pacific journal of cancer prevention: APJCP. 2018;19(6):1601–6.
15. Van Beekhuizen S, Ouwens MJ, Pham HA, Postma M, Heeg B. Pcn447 new methodology in parametric network meta analysis: non-mixture cure models. Value in Health. 2019;22:S523.
16. Amico M, Keilegom IV. Cure Models in Survival Analysis. Annual Review of Statistics and Its Application. 2018;5(1):311-42.
17. Othus M, Barlogie B, LeBlanc ML, Crowley JJ. Cure Models as a Useful Statistical Tool for Analyzing SurvivalCure Models and Multiple Myeloma. Clinical Cancer Research. 2012;18(14):3731-6.
18. Lambert PC, Thompson JR, Weston CL, Dickman PW. Estimating and modeling the cure fraction in population-based cancer survival analysis. Biostatistics. 2007;8(3):576-94.
19. Rahimzadeh M, Kavehie B, Zali MR. Cure models in analyzing long-term survivors. Translational Gastrointestinal Cancer. 2014;3(4):149-54.
20. Ghasemi F, Rasekhi A, Haghighat S. Analysis of the Survival of Breast Cancer Patients Using Weibull and Poisson Beta-Weibull Non-Mixture Cure Models. Research in Medicine: Journal of Research in Medical Sciences. 2019;42(4):236-42.
21. Rezaianzadeh A, Peacock J, Reidpath D, Talei A, Hosseini SV, Mehrabani D. Survival analysis of 1148 women diagnosed with breast cancer in Southern Iran. BMC cancer. 2009;9(1):1-11.
22. Nardin S, Mora E, Varughese FM, D'Avanzo F, Vachanaram AR, Rossi V, et al. Breast cancer survivorship, quality of life, and late toxicities. Frontiers in oncology. 2020;10.
23. Haque R, Ahmed SA, Inzhakova G, Shi J, Avila C, Polikoff J, et al. Impact of breast cancer subtypes and treatment on survival: an analysis spanning two decades. Cancer epidemiology, biomarkers & prevention. 2012;21(10):1848-55.
24. Usman M, Dikko H, Bala S, Gulumbe S. An application of Kaplan-Meier survival analysis using breast cancer data. Sub-Saharan African Journal of Medicine. 2014;1(3):132-7.