Journal of Basic Research in Medical Sciences

Journal HomePage : https://jbrms.medilam.ac.ir

Ilam University of Medical Sciences Publications

Risk Factors Associated with Overall Survival of Breast Cancer Patients Using Multivariate Cox Extended Models

Khadijeh Mamashli [1], **Hassan Nourmohammadi** [2], **Kourosh Sayehmiri** [1]✉

[1] Health faculty, Biostatistics Department, Ilam University of Medical Sciences, Ilam, Iran
[2] Medicine Faculty, Ilam University of Medical Sciences, Ilam, Iran

Article Info

Article type:
Original Article

Article History:
Received: Jul. 17, 2024
Revised: Jul. 19, 2025
Accepted: Aug. 08, 2025
Published Online: Sep. 30, 2025

✉ ***Correspondence to:***
Kourosh Sayehmiri
Health faculty, Biostatistics Department, Ilam University of Medical Sciences, Ilam, Iran

Email:
kourosh86@gmail.com

A B S T R A C T

Introduction: The present study aimed at identifying the risk factors associated with overall survival (OS) of breast cancer (BC) patients using multivariate Cox extended models.

Materials & Methods: This retrospective cohort study was conducted on 348 women with BC who were followed up for 10 years. Kaplan-Meier (KM) and log-rank statistics, Cox proportional hazard (PH), and multivariate Cox models were used to analyze the data. STATA V.17 and SPSS V.27 were used for data analysis.

Results: The median age of the patients was 55 years, and the median survival time was 29 months. Five- and 10-year OS were estimated at 93.4% and 88.4%, respectively. The results of multivariate analysis using the Cox model showed that lymph node (LN+) (hazard ratio (HR) = 2.86, P = 0.002), tumor size (HR = 1.99, P = 0.001), and progesterone receptor (PR-) (HR = 4.5, P = 0.002) increase death hazard significantly.

Conclusion: Prognostic factors indicated that women with lymph node involvement (LN+), positivity of human epidermal growth factor receptor 2 (HER2+), negativity of estrogen receptor (ER-), negative expression of progesterone receptor (PR-), advanced disease grade, and large tumor sizes were more likely to have a high hazard of death than other women.

Keywords: Breast Cancer, Overall survival, Cox extended models

➢ **Cite this paper**
Mamashli Kh, Nourmohammadi H, Sayehmiri K. Risk Factors Associated with Overall Survival of Breast Cancer Patients Using Multivariate Cox Extended Models. *J Bas Res Med Sci*. 2025; 12(4):12-19.

Introduction

According to the Worldwide Cancer Report, 6.8 million new cases of cancer and 2.4 million cancer-related deaths occurred in the world in 2018, out of which breast cancer (BC) accounted for 24.2% of all new cases and 15% of deaths (1). Generally, some predictors of the survival of BC patients have been reported to be age (2-4), tumor size (2, 5), lymph node involvement status (2, 6, 7), stage of disease (3, 8, 9), grade of disease (2), metastasis status (10), recurrence (10), genetic factors (11), and the type of treatment (4). The survival rate is the most important parameter for selecting cancer treatment and control strategies (12). Different statistical methods are used to identify the prognostic factors affecting cancer patients' survival (1). In this regard, the Cox PH model is a semi-parametric and popular model because it does not require a baseline hazard function for estimating the hazard ratio (HR) and regression coefficients, delivering results comparable to those of parametric models (13, 14). The key assumption of the Cox PH model is the proportionality of hazards for all predictive variables in the model over time (15). If the PH assumption is not met, the results of the Cox model may be misleading, whereas parametric models can be used as alternatives to overcome this problem during survival analysis. Most parametric models are based on accelerated failure time (AFT) models (16) and do not require the Cox PH assumption (13). The current study aimed at determining the predictors of OS in women with BC using the multivariate Cox PH models.

Materials and methods

Study Design

T This retrospective cohort study was conducted on 348 women with BC.

Setting and Participants

Clinic pathologic and survival data were obtained from Prat et al.'s study [GSE18229] (17), in which the patients were followed up for about ten years.

Measurements & Validity and Reliability

All the patients suffered from ductal-type BC, and variables such as age at diagnosis, grade of disease, tumor size, lymph node involvement status, and the expression status for estrogen receptor (ER), progesterone receptor (PR), and human epidermal growth factor receptor 2 (HER2) were reported as predictors of survival. Overall survival (in months), as the main outcome, was defined as the time interval from diagnosis to death.

Statistical and Data Analysis

The KM method was used to estimate the survival time function for all covariates, and the nonparametric log-rank test was applied to compare survival curves and determine the factors associated with OS. We used the goodness-of-fit [GOF] technique and a graphical method (a plot of ln(-ln S(t)) versus ln(t)) for different levels of variables to assess the PH assumption. The univariate Cox PH model was used to estimate the effects of covariates on OS. All covariates that were significant at the 0.1 level were entered into the multivariate Cox PH model (18). An extended Cox model was used for data analysis when the PH assumption was violated. In the presence of a categorical predictor variable, there is an increase in the proportion of the hazard changes in a continuous predictor, calculated from the Cox model (19). The hazard function graph was drawn against different Cox-Snell residuals in different models. The graph that was closer to the bisector line was considered the best model fitted to the data (20). Statistical analyses were performed using SPSS 27 and Stata 17.

Results

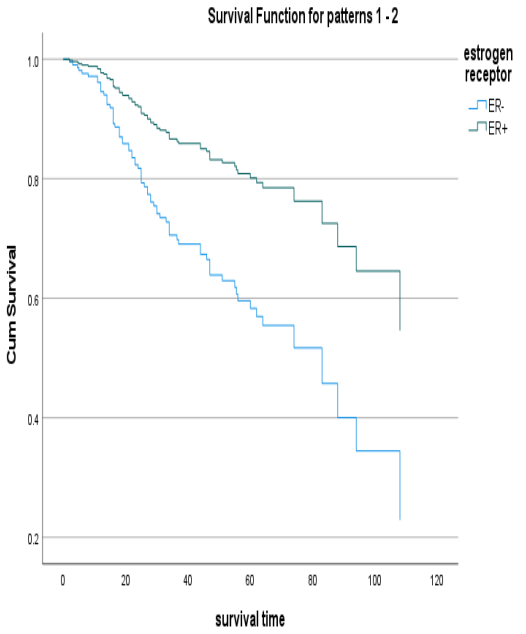
In this study, 348 women with BC were followed up for ten years after the diagnosis. The mean age of the patients at diagnosis was 57 years (range: 24-89 years). At the end of the follow-up period, 276 (79.3%) patients were alive, while 71 (20.7%) patients died. The mean survival time was 40 months, and the median survival time was 29 months (range:

1-118 months) (Table 1). Five- and 10-year OS were estimated at 93.4% and 88.4%, respectively. The results of the log-rank test showed a significant difference in the survival of patients with LN involvement (LN+) and LN- (P = 0.0002). Also, as the tumor size and disease grade increased, patients' survival decreased. There was a significant difference in the survival of patients regarding different disease grades (P=0.0054), HER2 status (P=0.0281), ER expression (P=0.0014), and PR expression (P=0.0004). The results of the log-rank test did not show a significant relationship between OS and age (P=0.66) (Table 1).

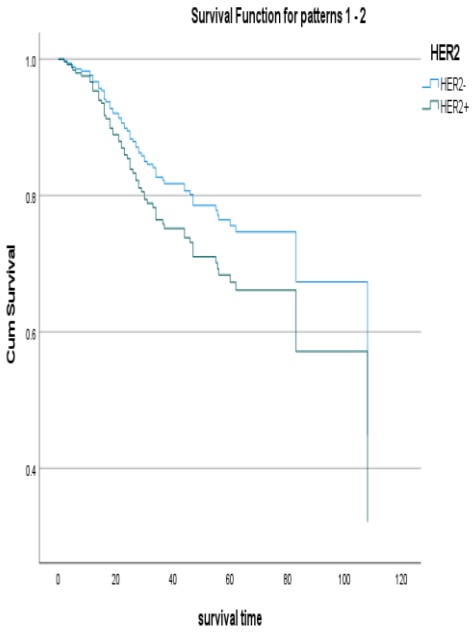
Table 1. Characteristics of patients and comparison of overall survival in terms of different categories of variables using the log-rank test.

Characteristics	N (%)			P (Log-rank test)
	Alive	death	Total	
Age <50 years =>50 years	121(81.2) 119 (76.7)	28(18.8) 36 (33.3)	149 (100) 169 (100)	0.66
LN status ln- Ln+	155(90.1) 121(69.9)	17 (9.9) 52 (30.1)	172 (100) 173(100)	0.0002
Her2 her2- her2+	184(81.8) 66(75)	41 (18.2) 22 (25)	225 (100) 88 (100)	0.0281
ER Er- Er+	94(67.6) 153(85.5)	45(32.4) 26(14.5)	139(100) 179(100)	0.0014
PR pr- pr+	97(67.8) 121(90.3)	46(32.2) 13(9.7)	143(100) 134(100)	0.0004
Grade G1 G2 G3	31(91.2) 100(87.7) 113(73.5)	3(8.8) 14(12.3) 48(26.5)	34(100) 114(100) 181(100)	0.0054
Tumor Size T1 T2 T3 T4	77(95.1) 151(83) 33(67.3) 13(41.9)	4(4.9) 31(17) 16(32.7) 18(58.1)	81(100) 182(100) 49(100) 31(100)	0.000

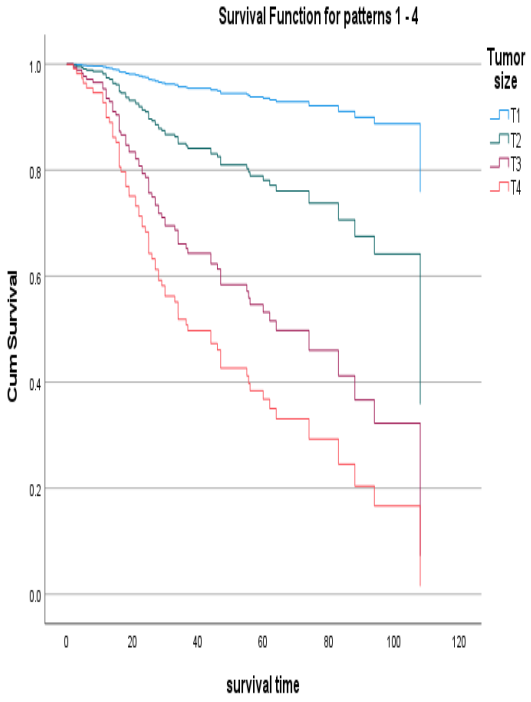
ER=Estrogen receptor, PR=Progesterone receptor, HER2=Human epidermal growth factor receptor2, LN= lymph node statues, Age= Age at diagnosis, Grade= Histological grade



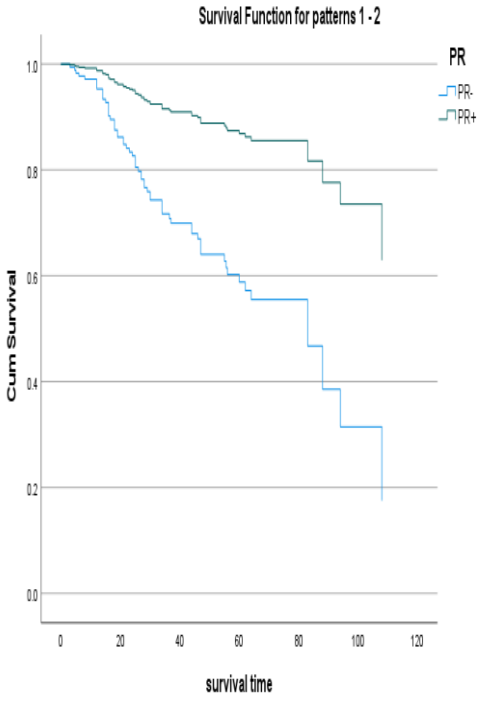
a



b



c



d

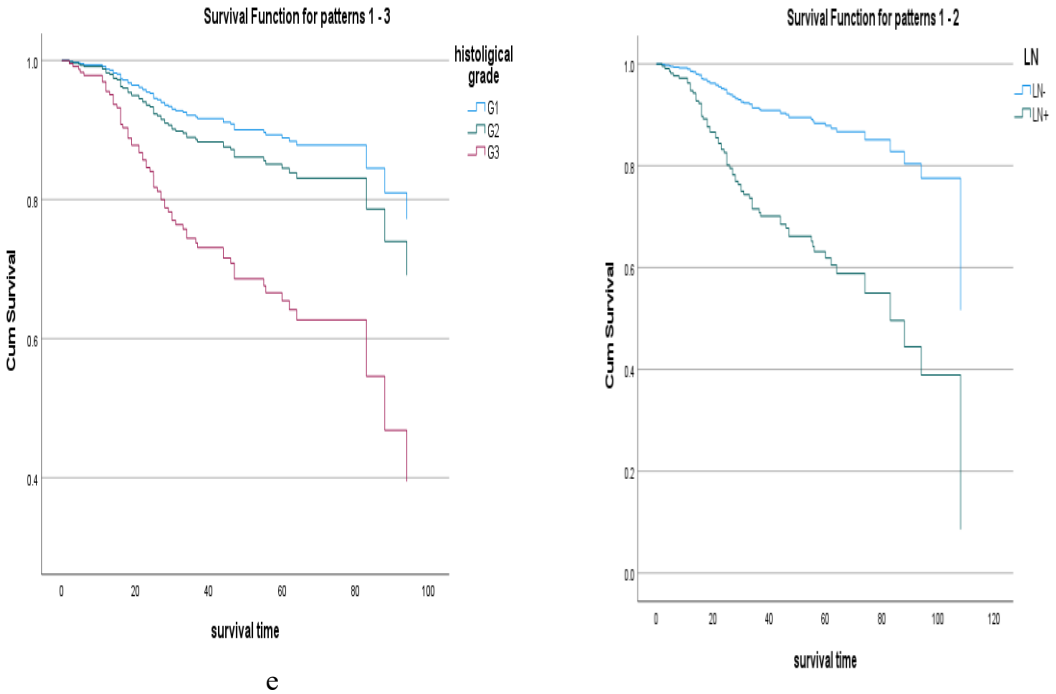


Figure 1. Comparison overall survival time in patients with breast cancer according to predictive factors, (a): ER,(b):HER2, (c);Tumor size ,(d):PR,(e):histological grade,(f):LN

Goodness-of-fit (GOF) and graphic methods showed that the PH assumption was met for all independent variables in the model, except for HER2 status (Table 2). For this variable, although the critical value was P = 0.109 (i.e., >0.05), the two survival curves were not parallel, suggesting the violation of the PH assumption (Fig. 1-b).

The findings of the Cox PH model in univariate analysis showed a statistically significant association between OS and each of LN involvement (P=0.000), HER2 status (P=0.032), PR expression (P=0.001), disease grade (G2, P=0.36; G3, P=0.038), and tumor size (T2, P=0.025; T3, P=0.003; and T4, P=0.000). However, there was no statistically significant association between age and OS (P = 0.67).

After checking all factors by a univariate model, variables retrieving P < 0.1 were entered into a multivariate Cox PH model. In the univariate model, estimated HR values for tumor size indicated that larger tumors elevated the risk of death. Since the PH assumption was violated regarding HER2 expression status, the Cox time-dependent model was employed to reduce the risk of bias.

Table 2. The results of the PH Cox model in univariate analysis for estimating overall survival in breast cancer patients and checking the fulfillment of the PH assumption.

Characteristics	Cox model (univariate)		PH assumption	P
	HR (95% CI)	P		
Age	0.89 (0.52-1.51)	0.67	Met	0.86
LN	2.95 (1.63-5.33)	0.00	Met	0.702
Positive HER2	1.98 (1.06-3.7)	0.032	Not met	0.109
ER	0.424 (0.25-0.73)	0.002	Met	0.294
PR	0.029 (0.142-0.61)	0.001	Met	0.054
Grade G1 G2	Ref 1.9 (0.45-8.96)	Ref 0.36	Met	0.091

G3	4.51 (1.08-18.8)	0.038		
Tumor size	Ref	Ref		0.093
T1	3.95 (1.19-13.1)	0.025	Met	
T2	6.6 (1.87-23.26)	0.003		
T3	14.1 (4.09-48.7)	0.00		
T4				

The results of the Cox extended model (Table 3) showed that the estimated HR for LN involvement was equal to 2.21, suggesting that women with LN+ were more than twice as likely to have lower OS than women with LN- (P = 0.038). As the 95% CI of HR for LN involvement did not cross 1, the effect of LN on survival was statistically significant. Although the effect of disease grade on survival was not statistically significant, the respective HR was greater than 1, indicating that the risk of death increased as BC grade increased (Table 3). The results of the extended model used have been presented in Table 3.

Table 3. The results of the multivariate Cox extended model for estimating the effects of different variables on overall survival in women with breast cancer.

Variables	Cox time-dependent coefficient	
	HR (95% CI)	P
LN (LN+ vs. LN-)	2.21(1.04-4.7)	0.038
HER2 (Her2- vs. Her2+)	0.61(0.14-2.64)	0.51
ER (Er+ vs. Er-)	0.91(0.35-2.32)	0.84
PR (Pr+ vs. Pr-)	0.4(0.12-1.27)	0.12
Grade G1	Ref	Ref
G2	1.17(0.13-10.66)	0.88
G3	3.13(0.37-26)	0.29
Size T1	Ref	Ref
T2	3.33(0.43-25)	0.25
T3	4.6(0.55-38.3)	0.15
T4	10.38(1.26-85.32)	0.029
T*Her2	-	0.075

Discussion

This study was conducted to determine the impacts of some prognostic factors on the OS of women with BC. We concluded that the Cox extended model was the best model to analyze the data and identify the prognostic factors associated with OS in women suffering from BC. Some factors, such as more advanced disease grade, larger tumor size, LN involvement, lack of PR expression, HER2 positive, and ER negative, were associated with increased risk of death in these patients.

The analysis utilizing a multivariate Cox extended model revealed distinct prognostic factors influencing overall survival (OS) in breast cancer patients. Notably, age at diagnosis demonstrated no statistically significant association with OS (HR=0.91, 95% CI: 0.35–2.32; P=0.67). This contrasts with some studies identifying age as a predictor (21, 22), but aligns with research suggesting age-related effects may be mediated by other variables, such as hormone receptor status (21, 23, 24). Conversely, estrogen receptor (ER) positive status emerged as a critical factor, consistent with established literature demonstrating improved

survival outcomes for ER+ tumors (21, 23). Tumor size had a significant association with survival (21, 23). In a meta-analysis done by Mulugeta, age at diagnosis, tumor size, and lymph node status were independent prognostic indicators for survival (23). The PH Cox model is one of the most commonly used models for the analysis of survival data, and its applicability has been approved in many studies (7, 25). The Cox extended and AFT models present a suitable description of survival data in most situations. Alfonso et al. (9) conducted a study on BC patients in Cuba, reporting GG as the best model and age and disease stage as the factors significantly associated with survival.

Limitations and Strengths

One limitation of this study is that we did not include treatment method; treatment method may be an independent variable on survival. The main focus of this study was to investigate the impact of some potential prognostic factors on the OS of BC patients. Further studies can be conducted to assess the impacts of different treatments on these parameters using different models. The log-logistic model is an appropriate model showing good fit into BC data, as reported by Amran et al. (4). In other studies (8, 10, 26), the Weibull model has been proposed as the most suitable model for analyzing BC data. In survival analysis we prefer to use simple models such as Cox PH models, but when the assumptions of these models are violated, we have to use Cox extended or AFT models.

Conclusion

In this study we showed when PH assumptions are violated for some risk factors, we have to use other models such as Cox extended models. Advanced disease grade, larger tumor size, LN involvement, lack of PR expression, HER2 positive, ER negative, and increased risk of death in patients.

Acknowledgements

We are grateful to Ilam University of Medical Sciences for approving this project.

Ethical Considerations

The Ethics Committee of the Ilam University of Medical Sciences approved the protocol of this research under the code of IR.MEDILAM.REC.1398.063.

Financial Disclosure

This research was a part of a M.S dissertation of biostatistics in Ilam University of Medical Sciences and supported by Ilam University of Medical Sciences.

Competing Interests' Disclosure

The authors declare that they have no competing interests.

Conflict of interest

The authors declare no conflicts of interests.

Authors' contributions

Conceptualization, Methodology, Validation, Formal Analysis, Investigation, Resources, Data Curation, Writing– Original Draft Preparation, Writing– Review & Editing, Visualization, Supervision, Software, Project Administration: KM, HN. KN.

Writing Disclosure

The writers attest that they worked alone to write and prepare this text, without the use of any professional writing services. Only the writers' original work and contributions are reflected in the text.

Data Availability Statement

The accompanying author may provide the data supporting the study's conclusions upon reasonable request.

References

1. Orbe J, Ferreira E, Núñez-Antón V. Comparing proportional hazards and accelerated failure time models for survival analysis. Stat Med. 2002;21(22):3493-510. https://doi.org/10.1002/sim.1251.
2. Elsayed M, Alhussini M, Basha A, Awad A. Analysis of loco-regional and distant recurrences in breast cancer after conservative surgery. World J Surg Oncol. 2016;14(1):144. https://doi.org/10.1186/s12957-016-0881-x.
3. Liu Z, Shen X, Liu R, Zhu G, Huang T, Xing M. Stage II Differentiated Thyroid Cancer Is a High-risk Disease in Patients <45/55 Years Old. J Clin Endocrinol Metab. 2019. https://doi.org/10.1210/jc.2018-02809.
4. Amran SE, Abdullah MAA, Kek SL, Jamil SAM. Analysis of survival in breast cancer patients by using different parametric models. J Phys Conf Ser. 2017; p. 012169.
5. Narod S. Tumour size predicts long-term survival among women with lymph node-positive breast cancer. Curr Oncol. 2012;19(5):249. https://doi.org/10.3747/co.19.1043.
6. Moghaddami FZ, Abolghasemi J, Asadi LM, Mohammadi M, Gohari M, Salehi M. Effective factors in the appearance of metastasis in patients with breast cancer using frailty model. J Arak Univ Med Sci. 2013;15(8):85-94.
7. Tonellotto F, Bergmann A, de Souza Abrahão K, de Aguiar SS, Bello MA, Thuler LCS. Impact of number of positive lymph nodes and lymph node ratio on survival of women with node-positive breast cancer. Eur J Breast Health. 2019;15(2):76. https://doi.org/10.5152/ejbh.2019.4414.
8. Baghestani A, Moghaddam S, Majd H, Akbari M, Nafissi N, Gohari K. Survival analysis of patients with breast cancer using Weibull parametric model. Asian Pac J Cancer Prev. 2015;16(18):8567-71. https://doi.org/10.7314/apjcp.2015.16.18.8567.
9. Alfonso AG, de Oca NAM. Application of hazard models for patients with breast cancer in Cuba. Int J Clin Exp Med. 2011;4(2):148.
10. Karimi A, Delpisheh A, Sayehmiri K. Application of accelerated failure time models for breast cancer patients' survival in Kurdistan Province of Iran. J Cancer Res Ther. 2016;12(3):1184.
11. Fallahzadeh H, Mohammadzadeh M, Pahlavani V, Pahlavani N. A study on the prognostic factors of breast cancer survival time using Bayesian Cox model. J Isfahan Med Sch. 2018;36(466):49-55. https://doi.org/10.22122/JIMS.V36I466.9108.
12. Ederer F. The relative survival rate: a statistical methodology. J Natl Cancer Inst. 1961;6:101-21.
13. Kleinbaum DG, Klein M. The Cox proportional hazards model and its characteristics. Survival analysis: Springer; 2012. p. 97-159.
14. Khoei R, Bakhshi E, Azarkeivan A, Biglarian A. Survival analysis of the thalassemia major patients using parametric and semiparametric survival models. JHA. 2015;18(59).
15. Klein JP, Moeschberger ML. Survival analysis: techniques for censored and truncated data. Springer Science & Business Media; 2006.
16. Montaseri M, Charati JY, Espahbodi F. Application of parametric models to a survival analysis of hemodialysis patients. Nephrourol Mon. 2016;8(6). https://doi.org/10.5812/numonthly.28738.
17. Prat A, Parker JS, Karginova O, Fan C, Livasy C, Herschkowitz JI, et al. Phenotypic and molecular characterization of the claudin-low intrinsic subtype of breast cancer. Breast Cancer Res. 2010;12(5):2. https://doi.org/10.1186/bcr2635.
18. Qi J. Comparison of proportional hazards and accelerated failure time models. 2009.
19. George B, Seals S, Aban I. Survival analysis and regression models. J Nucl Cardiol. 2014;21(4):686-94.
20. Hosseini M, Mohammad K, Rahimzadeh M, Mahmoodi M. Comparison of survival models in studying breastfeeding duration. Hakim J. 2007;10(1):66-71.
21. Mulugeta C, Emagneneh T, Kumie G, Sisay A, Abebaw N, Ayele M, et al. Predictors of survival rates among breast cancer patients in Ethiopia: a systematic review and meta-analysis 2024. Arch Public Health. 2025;83(1):30. https://doi.org/10.1186/s13690-025-01514-6.
22. Abedi G, Janbabai G, Moosazadeh M, Farshidi F, Amiri M, Khosravi A. Survival rate of breast cancer in Iran: a meta-analysis. Asian Pac J Cancer Prev. 2016;17(10):4615-21. https://doi.org/10.22034/apjcp.2016.17.10.4615.
23. Jayasinghe UW, Taylor R, Boyages J. Is age at diagnosis an independent prognostic factor for survival following breast cancer? ANZ J Surg. 2005;75(9):762-7. https://doi.org/10.1111/j.1445-2197.2005.03515.x.
24. Lees AW, Jenkins HJ, May CL, Cherian G, Lam EW, Hanson J. Risk factors and 10-year breast cancer survival in northern Alberta. Breast Cancer Res Treat. 1989;13(2):143-51. https://doi.org/10.1007/bf01806526.
25. Abadi A, Yavari P, Dehghani-Arani M, Alavi-Majd H, Ghasemi E, Amanpour F, et al. Cox models survival analysis based on breast cancer treatments. Iran J Cancer Prev. 2014;7(3):124.
26. Ghorbani N, Yazdani-Charati J, Anvari K, Ghorbanifar N. Application of Weibull accelerated failure time model on the disease-free survival rate of breast cancer. Iran J Cancer Prev. 2016;4(2):11-8. https://doi.org/10.18869/acadpub.jhs.4.2.11.