International Review of Applied Sciences
Vol. 11, No. 1, 32-39, 2025
ISSN(E) 2411-667X/ ISSN(P) 2518-0207
DOI: 10.20448/iras.v11i1.7770

Modelling the survival rate of breast cancer patients: Non-parametric and semi-parametric approach

Ayinla, Faruq Ayomide[1]✉
Adejumo, Adebowale Olusola[2]

[1,2]*Department of Statistics, Faculty of Physical Sciences, University of Ilorin, Nigeria.*
[1]*Email: faruqayinla80@gmail.com*
[2]*Email: aodejumo@unilorin.edu.ng*


(✉ Corresponding Author)

Abstract

The comparison of survival among patients with breast cancer was analyzed using the Cox proportional hazards model and the Kaplan-Meier estimator to model the survival outcomes. A sample of 563 patients was selected to evaluate the impact of risk factors such as age, tumor stage, type of surgery, chemotherapy, radiotherapy, and hormone therapy. The log-rank test results indicated that radiotherapy ($p = 0.8$) and age ($p = 0.08$) did not significantly affect survival distribution. Conversely, tumor stage ($p < 0.0001$), surgical type ($p = 0.0004$), chemotherapy ($p = 0.04$), and hormone therapy ($p = 0.0004$) showed significant differences in survival outcomes. Cox regression analysis identified tumor stage as the most influential factor, with patients at Stage 2, 3, and 4 exhibiting 1.85 to 3.98 times higher risks of mortality compared to those at Stage 1. Additionally, patients undergoing mastectomy had a 1.32 times higher risk than those receiving breast-conserving surgery ($p = 0.0288$). The assumptions of proportional hazards were violated for chemotherapy and hormone therapy ($p = 0.015$ and $p = 0.018$, respectively), which contributed to the overall significance of the model ($p = 0.048$). Stratification of these variables confirmed the robustness of the model, with tumor stage and surgical option remaining key factors. These findings underscore the importance of early diagnosis in improving survival outcomes for breast cancer patients.

Keywords: Breast cancer, Cox proportional hazard, Kaplan-Meier, Survival rate, Stratification, Hazard Ratio, Early diagnosis.

Citation | Ayomide, A. F., & Olusola, A. A. (2025). Modelling the survival rate of breast cancer patients: Non-parametric and semi-parametric approach. *International Review of Applied Sciences*, *11*(1), 31–39. 10.20448/iras.v11i1.7770.
History:
Received: 22 September 2025
Revised: 9 October 2025
Accepted: 11 November 2025
Published: 28 November 2025

Publisher: Asian Online Journal Publishing Group

Funding: This study received no specific financial support.
Institutional Review Board Statement: Not applicable.
Transparency: The authors confirm that the manuscript is an honest, accurate, and transparent account of the study; that no vital features of the study have been omitted; and that any discrepancies from the study as planned have been explained. This study followed all ethical practices during writing.
Competing Interests: The authors declare that they have no competing interests.
Authors' Contributions: Both authors contributed equally to the conception and design of the study. Both authors have read and agreed to the published version of the manuscript.

Contents

1. Introduction **33**
2. Methodology **34**
3. Results and Discussion **35**
4. Conclusion **39**
References **39**

Contribution of this paper to the literature

This study contributes to the existing literature by incorporating both non-parametric and semi-parametric models in breast cancer prognosis. The main contribution is the tumor stage and the surgery type used as the key elements of patient survival. Stratification improves model validity, and individualized and early treatment improve outcomes.

1. Introduction

Breast cancer is among the most common or prevalent diseases that women in virtually every part of the world have in common, and this has always been a major challenge to the health of the population over the last few decades, notwithstanding the early detection, mode of treatment, and even the survival of patients with breast cancer [1]. There has been a phenomenal change in the survival of breast cancer patients, and the data now clearly demonstrate a significant improvement in both overall survival rates and disease-specific survival rates of breast cancer patients. The awareness of these survival patterns is highly relevant to clinicians, researchers, and patients, as it aids in making informed treatment choices and addressing prognostic challenges within the contemporary oncological environment [2]. The recent trends in survival rates provide an optimistic future outlook concerning breast cancer survival rates, as women who receive a breast cancer diagnosis collectively have a 5-year relative survival rate of nearly 91 percent [3]. This progress is an effect of the overall positive outcomes of more efficient screening activities, earlier identification measures, as well as the design of more effective treatment measures [4]. Survival rates have also allowed us to determine that early detection is incredibly important, with only the localized phase of breast cancer being highly positive, as the relative survival rate at that stage of the disease at the time of detection is 99.6% after five years [5]. Regional disease (growth to surrounding structures and lymph nodes) exhibits a considerably lower relative survival rate at 86% after five years. Conversely, metastatic disease, which is far more advanced, has a more challenging prognosis, as women have a 31% relative survival at five years in reference to the continued investigation necessary in expanded illness care [1].

The temporal trends in the mortality of breast cancer further support the suggestion of developments in patient outcomes. The death due to breast cancer has decreased significantly since 1989, when the rates were recorded as high, with a total of 44 percent reduction in breast cancer mortality rates, and this figure translates to almost half a million women being saved due to the lack of breast cancer-related deaths in the United States [4]. This outstanding success is due to improvements in the effectiveness of treatment and earlier diagnosis. Surprisingly, although mortality rates indicate a downward trend, the number of new female breast cancer cases has been increasing on average by 1.0 percent every year over the period between 2012 and 2021, which is an indicator of how well modern treatment works despite its rising incidence rates [2]. The heterogeneity of breast cancer survival is not limited to the stage at diagnosis; it also involves biological subtypes, demographic factors, and treatment response patterns. The statistics show that 63.5 percent of breast cancers are diagnosed at the local level as a result of the effectiveness of screening programs and awareness [5]. Nonetheless, disparities in survival prevail among various groups in the population and between tumor attributes, and further investigation into the factors affecting survival rates and the implementation of individual treatments is still needed [6]. The survival statistics can be studied not only as evidence of the improvement that occurred in breast cancer care but also as a basis for the studies conducted to enhance patient outcomes further. As we become more aware of breast cancer biology, develop new therapeutic agents, and determine the best use of current and new agents, the analysis of patterns of survival becomes much more detailed, necessitating increasingly sophisticated statistical methods to model the patterns of factors affecting patient prognosis.

The temporal evolution of breast cancer mortality numbers supports the indication of improved patient outcomes. The burden of death due to breast cancer has decreased significantly since it was recorded as high in 1989, when a 44 percent reduction in breast cancer death rates was observed. This reduction has translated into approximately half a million women saved in the United States due to a decline in breast cancer-related deaths [4]. This remarkable achievement is attributed to increases in the accuracy of treatment as well as early diagnosis. Surprisingly, the mortality rates show a statistically decreasing trend, but the new cases of breast cancer in women are, on average, increasing at an annual rate of 1.0 percent across the years between 2012 and 2021, which points towards the efficiency of modern treatment that is associated with its increasing cases [2]. The diversity of breast cancer survival extends not only to the stage at which the cancer is diagnosed but also includes biological subtypes, demographic factors, and patterns of response to treatment. According to statistics, 63.5 percent of breast cancers are diagnosed at the local level due to the efficiency of screening programs and awareness [5]. However, a survival gap remains between different groups within the population and between characteristics of the tumor, and further research on the factors that influence the survival rate and the introduction of personalized treatments is still required [6].

The survival statistics may be reviewed as a sign of the progress taking place regarding breast cancer care, but may also be discussed as a foundation of the research performed to improve the situation and outcomes of patients even more. With increasing knowledge of the biology of breast cancer, new agents to treat it, and understanding of how to harness current and novel agents, the description of survival patterns will be much more detailed and will require increasingly complex statistical methods to describe the patterns of factors impacting patient prognosis.

Cancer survivorship has since shifted to a more holistic approach, extending beyond survival thresholds to a multi-dimensional focus on the physical, emotional, and social well-being of individuals during the cancer experience [7]. Such a comprehensive measurement of the outcome requires advanced analytical procedures capable of considering competing risks, covariates that change over time, and patient-reported outcomes [8].

The increasing use of precision medicine strategies has added a new layer of complexity to survival analysis because treatment decisions are becoming more tailored to individual patient and tumor characteristics. Oncotype DX, MammaPrint, and other genomic profiling tools have provided practical prognostic and predictive data, which contribute to treatment choices and, by extension, survival rates [9, 10]. Inclusion of these molecular markers in

survival models involves complex statistical techniques that may involve large dimensions of information and intricate interaction effects [11].

The study investigates the impact of visual analytics on understanding breast cancer trends, patient demographics, clinical characteristics, and treatment outcomes, facilitating better decision-making in healthcare. The objectives are to: examine the relationship between patients' demographic and clinical characteristics (such as age, tumor stage, and surgery type) and their survival outcomes; also to assess the impact of different breast cancer treatment modalities (chemotherapy, hormone therapy, and radiotherapy) on patient survival rates; and to evaluate the statistical significance of survival differences across various treatment groups using chi-square tests (Log-rank tests) and Kaplan-Meier survival analysis.

2. Methodology

A log-rank test, which is one of the most used statistical tests in the analysis of survival, was first proposed by Mantel [12] and later refined by Peto and Peto [13]. It is a non-parametric test tailored specifically to compare survival distributions across two or more groups and thus invaluable in clinical trials, epidemiological research, and other studies where time-to-event outcomes are central to inquiry.

The popularity of the test is based on its capability to effectively analyze censored data and assumes very few distributional assumptions. In contrast to parametric tests, which need to specify the underlying survival distribution, the log-rank test can be performed within a non-parametric model, providing relatively robust information under a broad spectrum of survival patterns.

The underlying idea of the log-rank test is that the observed and expected numbers of events are compared within each category under the null hypothesis of the same survival distributions. Comparison is done at each observed event time, and some information on the entire follow-up time is contained within the test statistic. The statistic obtained follows a chi-square distribution, which facilitates statistical inference on the null hypothesis.

For comparing k groups, the log-rank test addresses:

$S_1(t) = S_2(t) = ... = S_k(t)$ for all $t \geq 0$ against.

$S_i(t) \neq S_j(t)$ for at least one pair (i,j) and some $t \geq 0$.

The null hypothesis states that all groups have identical survival distributions, while the alternative hypothesis allows for differences in survival between at least two groups.

Under the assumption of proportional hazards, the equivalent is to:

$h_1(t) = h_2(t) = ... = h_k(t)$ for all $t \geq 0$.

This equivalence is important because it connects the log-rank test to the broader framework of proportional hazards models.

Consider two groups with survival times and censoring indicators:

Group 1: $(Y_{11}, \delta_{11}), (Y_{12}, \delta_{12}), ..., (Y_{1n1}, \delta_{1n1})$.

Group 2: $(Y_{21}, \delta_{21}), (Y_{22}, \delta_{22}), ..., (Y_{2n2}, \delta_{2n2})$.

Let $t_1 < t_2 < ... < t_a$ be the D distinct event times across both groups.

At each event time t_i, we construct a 2×2 contingency table.

Table 1. Summary of events and number at risk for two groups in survival analysis.

Group	Events	At Risk	Total
Group 1	d_{1i}	n_{1i}	n_{1i}
Group 2	d_{2i}	n_{2i}	n_{2i}
Total	d_i	n_i	n_i

Table 1 presents the number of events (e.g, death or failure) and the number of people at risk at each observed time in two groups in a survival study.

Where:

d_{1i}, d_{2i} = Number of events in groups 1 and 2 at time t_i.

n_{1i}, n_{2i} = Number at risk in groups 1 and 2 just before time t_i.

$d_i = d_{1i} + d_{2i}$ = Total events at time t_i.

$n_i = n_{1i} + n_{2i}$ = Total at risk just before time t_i.

The most commonly used multivariate method for testing survival data in medical research is the Cox Proportional Hazards model, developed by Cox [14]. This semi-parametric regression model establishes the relationship between event occurrence rates and multiple covariates through the hazard function, which represents the instantaneous probability of an event occurring at any given time point. Mathematically expressed as.

$$h(t) = h_o(t) exp^{\beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 + \cdots + \beta_p X_p} \quad (1)$$

The model consists of a baseline hazard function $h_0(t)$ and a linear combination of p covariates (x_1 denoting the age category, x_2 denoting the tumor stage, x_3 denoting the surgery category, x_4 denoting chemotherapy, x_5 denoting radiotherapy, and x_6 denoting hormone therapy) with their corresponding coefficients ($\beta_1, \beta_2, ..., \beta_p$). The main benefit of this method is that the baseline hazard is estimated using non-parametric methods, which do not require that the survival times have a particular statistical distribution, unlike in the conventional parametric survival models.

This model is a multiple linear regression of the logarithm of the hazard on the covariates, in which the baseline hazard is a time-varying intercept term. The covariates are multiplicative on the hazard at any given time, and the basic proportional hazards assumption is that the hazard ratio between any two groups does not change with time. Consequently, the hazard curves of different groups should be proportional and should never intersect. The exponential changes of the regression coefficients, exp (β_i), are hazard ratios, and when they are not equal to one, they are covariates associated with the probability of events (negative with the length of survival) or covariates that are protective. Despite the fact that this assumption of proportional hazards is usually quite

reasonable in the case of survival data, it should be strictly examined and confirmed to guarantee valid interpretation of the model and sound conclusions.

3. Results and Discussion

The paper includes a number of statistical methods to evaluate the correlation between patient features and survival rates. In particular, tests of association between the dependent variable (Status: Alive/Dead) and different independent variables were tested using the Log-rank test. Additionally, survival analysis methods such as the Kaplan-Meier (KM) estimator and Cox proportional hazards model are employed to further study how patient survival depends on other factors. The Kaplan-Meier technique is applied to obtain estimates of survival functions for various subgroups of the sample, providing a graphical illustration of survival probabilities over time. The Cox proportional hazards model is used to evaluate the effects of continuous and categorical variables on the risk of mortality, controlling for other possible confounders. These techniques offer a comprehensive understanding of the variables affecting patient outcomes, highlighting important predictors of survival and mortality risk.

Table 2 offers a descriptive analysis of different clinical and tumor features using a sample of 1,090 observations that was subsequently narrowed to 563 observations due to the consequence of survival time that is too consequential and was subsequently capped at a maximum of 10 years. Kaggle is the source of the dataset that provides extensive and detailed clinical data that can be utilized in research. The link to the dataset is available here: https://www.kaggle.com/code/alexandervc/breast-cancer-metabric-with-omics-data-1

According to the statistical analysis of this cancer data (n=563 patients), the cohort is highly clinically diverse in all the measured parameters. The mean age of the patients at the time of diagnosis is 62.28 (SD=13.83), but there is a slightly negative skewness (-0.31), which demonstrates the presence of a slight concentration of patients in their sixties and fewer younger ones. The median age of 63 years supports the typical diagnosis at the beginning of the sixties, with ages ranging from 22 to 96 years. The overall survival varies significantly, with a mean of 63.05 months (SD=32.45) and a slightly normal distribution (skewness=0.04), although the survival range is broad, from 0.1 to 119.87 months, indicating heterogeneity among patients. The average relapse-free survival is 52.68 months (SD=34.26) with positive skewness (0.35), suggesting that most patients experience shorter disease-free survival, with a longer tail representing those in sustained remission. Tumor characteristics exhibit considerable heterogeneity; the average tumor size is 28.52 cm (SD=17.28), with a highly skewed distribution (skewness=3.57) and kurtosis of 21.54. This pattern indicates that most patients have smaller tumors, while a significant number have very large masses, ranging from 1 to 180 cm. Such a wide variation in clinical parameters reflects the multifaceted nature of this malignancy, with early mortality (minimum 0.1 months) and prolonged survival (maximum 119.87 months) as evidenced by survival and relapse-free survival data. The data confirms that older adults are generally more susceptible to this cancer and highlights the broad spectrum of disease presentation and progression, emphasizing the challenge in predicting individual patient outcomes within this diverse group.

Table 2. Descriptive summary table.

Variable	Age at diagnosis	Overall survival (Months)	Relapse-free status (Months)	Tumor size
Mean	62.284	63.054	52.675	28.515
Standard error	0.583	1.368	1.444	0.728
Median	63	61.9	45.933	25
Mode	69	19.733	37.467	20
Standard deviation	13.832	32.453	34.256	17.276
Sample variance	191.311	1053.205	1173.440	298.457
Kurtosis	-0.498	-1.108	-1.120	21.543
Skewness	-0.309	0.0389	0.351	3.566
Range	74	119.767	119.733	179
Minimum	22	0.1	0	1
Maximum	96	119.867	119.733	180
Sum	35066	35499.633	29656.067	16054
Count	563	563	563	563

Table 3. Frequency table presenting the variables and the test of independence between survival status and various independent variables.

Risk Factors		Event (Death)	Censored (Alive)	Total	Log rank value	P value
Age category	Young	52 (72%)	20 (28%)	72	6.6	0.08
	Middle Age	99 (61%)	63 (39%)	162		
	Elderly	140 (61%)	48 (39%)	188		
	Senile	87 (86%)	14 (14%)	101		
Tumor stage	Stage I	94 (66%)	49 (34%)	143	225	<0.0001
	Stage II	263 (76%)	85 (24%)	348		
	Stage III	56 (84%)	11 (16%)	67		
	Stage IV	5 (100%)	0 (0%)	5		
Surgery type	Breast conserving	138 (66%)	71 (34%)	209	12.6	0.0004
	Mastectomy	280 (79%)	74 (21%)	354		
Chemotherapy	Yes	112 (72%)	43 (28%)	155	4.1	0.04
	No	306 (75%)	102 (25%)	408		
Radiotherapy	Yes	265 (72%)	105 (28%)	370	0.1	0.8
	No	153 (79%)	40 (21%)	193		
Hormone therapy	Yes	261 (71%)	108 (29%)	369	12.3	0.0004
	No	157 (81%)	37 (19%)	194		
	Total	**418 (74%)**	**145 (26%)**	**563**		

Table 3 displayed how the risk factors impact survival outcomes, survival analysis, and it can be observed that tumor stage, surgery type, chemotherapy, and hormone therapy all significantly impact patient outcomes, showing clear differences in survival distributions. The log-rank test confirms that these variables have a meaningful effect on survival, with p-values below 0.05 indicating statistical significance. Among these tumors, stage is the strongest predictor of survival ($p < 0.0001$). Older patients and those diagnosed at advanced tumor stages (III & IV) have considerably lower survival rates (11% and 0%), reinforcing the importance of early detection and timely intervention to improve survival chances.

Surgery is another factor that has been found to be more successful in terms of survival than mastectomy surgery, which can be explained by the fact that patients who are more severely ill might have surgery as opposed to those who do not. Chemotherapy ($p = 0.04$) and radiotherapy ($p = 0.0004$) are effective methods of treatment that increase survival, but their efficacy depends on the specifics of a patient and on the progression of the disease.

Frequency distribution-wise, the data indicate that a high percentage of the patients are within the middle-aged and elderly groups, with higher mortality rates as age advances. Stage II tumors represent the greatest proportion of cases, but survival rates decrease drastically as the disease progresses, thus indicating the aggressiveness of late-stage cancer. The allocation of treatment modalities indicates that chemotherapy, radiotherapy, and hormone therapy are extensively utilized, but their effects on survival vary among different patient groups.

Influence of Different Breast Cancer Factors on Patient Survival Outcomes?

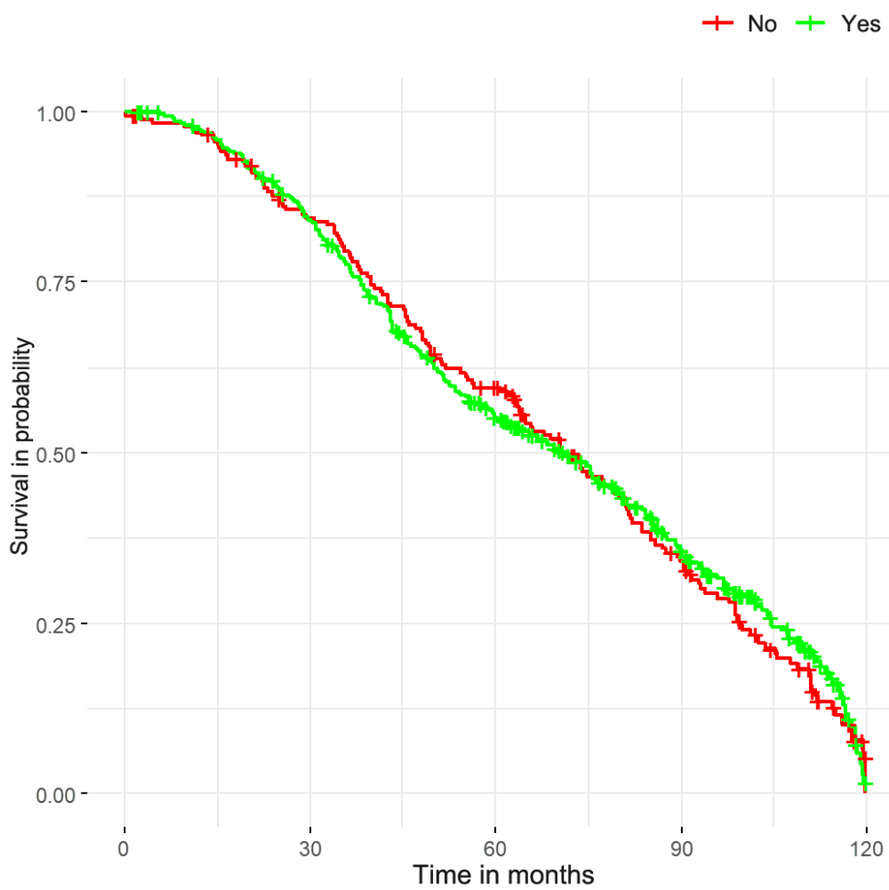

Figure 1. Kaplan-Meier survival curve by radiotherapy status.

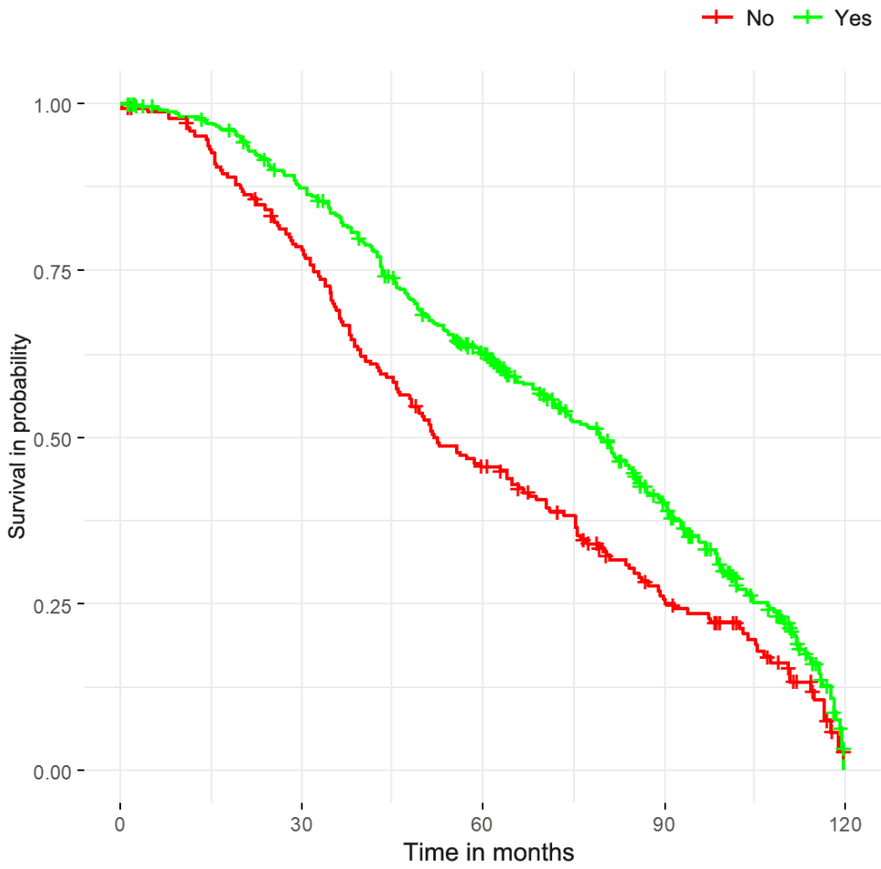

Figure 2. Kaplan-Meier survival curve by hormone therapy status.

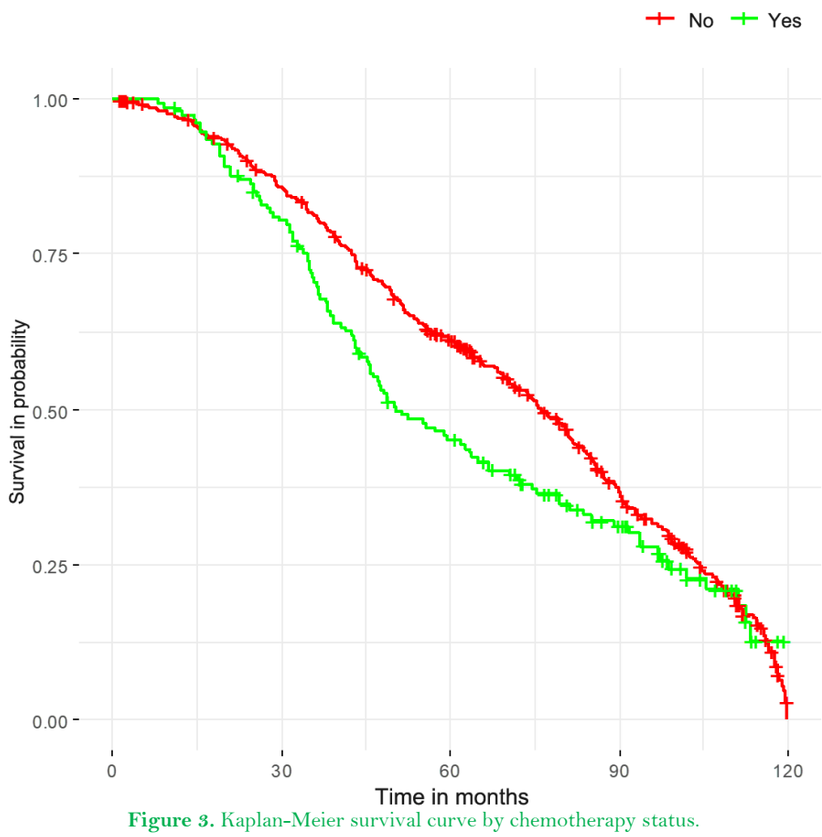

Figure 3. Kaplan-Meier survival curve by chemotherapy status.

Figure 4 illustrates the survival probability of both the surgical groups. Individuals with mastectomy enjoyed better survival than those in the group of breast-conserving surgery in the follow-up. This difference in survival between the two groups is higher in the period beyond the first 40 months, with the breast-conserving group experiencing a greater decline in the likelihood of survival. The mastectomy group had a higher survival rate within the period of observation, which indicates that the long-term outcome is more favorable. The difference in the curves proves that the survival of breast-conserving surgery is worse, and the factor is not temporary but observed throughout the entire period of 120 months.

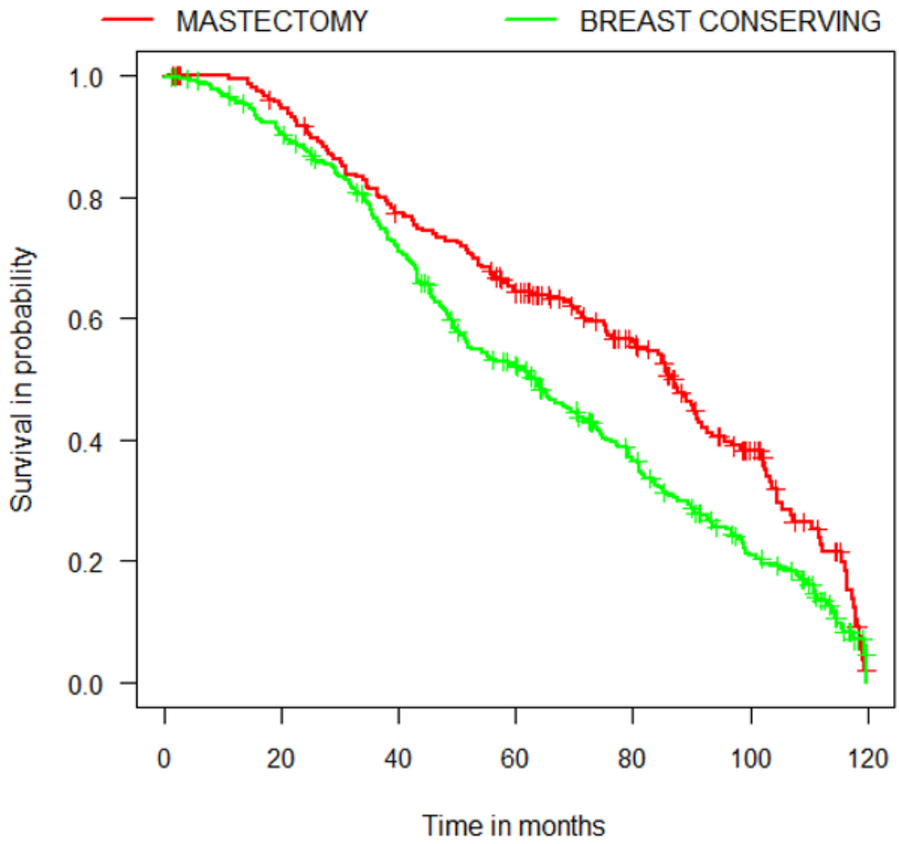

Figure 4. Kaplan-Meier survival curve by surgery type status.

The radiotherapy analysis (Figure 1) shows that there was no statistically significant difference in the survival probability of the patients undergoing radiotherapy and those who did not receive radiotherapy, as the two curves share a near-identical trajectory over the course of observation. This observation indicates that radiotherapy might not have a significant survival benefit for this particular group of patients. On the other hand, the survival differences are also significant between both hormonotherapy (Figure 2) and chemotherapy (Figure 3), with treated individuals recording significantly better survival rates than the untreated controls. The benefit of hormonotherapy is most significant during the intermediate period of follow-up (Months 30-60), when a significant separation between the treatment groups is also observed. The comparison of chemotherapy shows a slower but longer-term survival response, whereby the treatment advantage becomes more pronounced after the first 24 months of treatment, and the response continues throughout the follow-up period.

Null Hypothesis (H_0): The hazards are proportional (no violation of the PH assumption).
Alternative Hypothesis (H_1) : The hazards are not proportional (violation of the PH assumption).

Table 4. Testing of the Validity of the Factors.

Risk factors	Chi-Square (χ^2)	df	P
Chemotherapy	5.921	1	0.015
Radiotherapy	0.254	1	0.614
Hormone Therapy	5.624	1	0.018
Age	6.214	3	0.102
Tumor Stage	6.724	3	0.081
Surgery Type	0.629	1	0.428
GLOBAL	**18.462**	**10**	**0.048**

Based on Table 4, it is clear that out of the six factors under analysis, only Chemotherapy and Hormone Therapy do not uphold the proportional hazards assumption, with p-values of 0.015 and 0.018, respectively. These values are considered statistically significant and result in the rejection of the null hypothesis that states that the hazard ratios do not change with time. The model, therefore, has to be adjusted by stratifying these two variables. Their breach also contributes to the overall breach of the global proportional hazards assumption, which also warrants adjustment.

Table 5. Testing of Validity of the Factors After Stratification.

Variable	Chi-Square (χ^2)	Df	p-value
Radiotherapy	0.0833	1	0.77
Age	1.6713	3	0.64
Tumor stage	3.5444	3	0.32
Surgery type	0.4373	1	0.51

Table 5 presents the report after stratification of the two factors that violated the proportional hazards (PH) assumption, Chemotherapy and Hormone Therapy. As noted, the rest of the variables met the assumption, and

their p-values were higher than the 0.05 level of significance. This indicates that the proportionality of hazards holds for these variables. Therefore, it can be concluded that Radiotherapy, Age, Tumor Stage, and Surgery Type are valid predictors and significantly contribute to the survival outcomes of breast cancer patients within the context of the Cox proportional hazards model.

Table 6. Cox proportional Hazards model results.

Risk factors	Coef	Exp (Coef)	SE (Coef)	Z	P-value
Treatments (Reference: No)					
Radiotherapy	0.052	1.054	0.126	0.415	0.678
Age category (Reference: Elderly)					
Young	0.056	1.057	0.189	0.295	0.768
Middle Age	-0.047	0.954	0.143	-0.329	0.742
Senile	0.211	1.235	0.135	1.568	0.117
Tumor stage (Reference: Stage 1)					
Stage 2	0.617	1.853	0.137	4.507	<0.0001
Stage 3	0.977	2.657	0.209	4.682	<0.0001
Stage 4	1.381	3.981	0.475	2.909	0.004
Surgery type (Reference: Breast conserving)					
Mastectomy	0.278	1.320	0.127	2.186	0.029

$$h(t) = h_0 e^{0.6166\ \text{stage 2} + 0.9770\ \text{stage 3} + 1.3814\ \text{stage 4} + 0.2776\ \text{mastectomy}} \quad (2)$$

Table 6 presents the results of the logistic regression analysis. It reveals that the factors bearing significance in the outcome are the tumor stage and surgery type, whereas treatment type and age category do not. The outcome was increasingly more pronounced with patients with higher levels of disease, as Stage 1 (OR = 1.79, p = 0.0001), compared to Stage 2 (OR = 1.85, p = 0.0001), Stage 3 (OR = 2.66, p = 0.0001), and Stage 4 (OR = 3.98, p = 0.0036), which are significantly pronounced. Additionally, a mastectomy was found to lead to a 32 percent increase in the chances of the outcome compared to breast-conserving surgery (p = 0.0288). Conversely, receiving radiotherapy (OR = 1.05, p = 0.678) and being in different age categories (young, middle-aged, senile) did not have any statistically significant impact on the outcome. This indicates that the stage of the tumor and the type of operation are more important factors in determining the likelihood of the outcome in this model, compared to age and radiotherapy.

4. Conclusion

The hormonotherapy benefit is most pronounced during the intermediate follow-up period (months 30-60), where a substantial separation between treatment groups is observed. The comparison of chemotherapy indicates a slower, yet significant progress of survival rates, and the benefit of the treatment can be seen after the first 2 years of therapy and continues during the follow-up period. The paper concluded that radiotherapy, age, tumor stage, and surgery type are valid predictors and significantly contribute to the survival outcomes of breast cancer patients within the context of the Cox proportional hazards model. Through logistic regression, it was also concluded that tumor stage and surgical option are the key factors of breast cancer survival, and the necessity of early diagnosis and treatment based on an individual approach.

References

[1] American Cancer Society, *Breast cancer facts & figures 2024–2025*. Atlanta: American Cancer Society, 2024.

[2] J. S. Siegel *et al.*, "Psilocybin desynchronizes the human brain," *Nature*, vol. 632, no. 8023, pp. 131-138, 2024. https://doi.org/10.1038/s41586-024-07624-5

[3] Centers for Disease Control and Prevention, *U.S. cancer statistics — female breast cancer stat bite*. United States: Centers for Disease Control and Prevention, 2025.

[4] A. N. Giaquinto *et al.*, "Breast cancer statistics 2024," *CA: A Cancer Journal for Clinicians*, vol. 74, no. 6, pp. 477-495, 2024. https://doi.org/10.3322/caac.21863

[5] National Cancer Institute, *Cancer stat facts: Female breast cancer — surveillance, Epidemiology, and End Results (SEER) program*. United States: U.S. National Institutes of Health, 2024.

[6] C. E. DeSantis *et al.*, "Breast cancer statistics, 2019," *CA: A Cancer Journal for Clinicians*, vol. 69, no. 6, pp. 438–451, 2019.

[7] P. L. Hewitt, G. L. Flett, S. B. Sherry, and C. Caelian, "Trait perfectionism dimensions and suicidal behavior. In T. E. Ellis (Ed.), Cognition and suicide: Theory, research, and therapy." Washington, DC: American Psychological Association, 2006, pp. 215–235.

[8] P. C. Austin, A. Latouche, and J. P. Fine, "A review of the use of time-varying covariates in the Fine-Gray subdistribution hazard competing risks regression model," *Statistics in Medicine*, vol. 39, no. 3, pp. 103–113, 2019.

[9] S. Paik *et al.*, "A multigene assay to predict recurrence of tamoxifen-treated, node-negative breast cancer," *New England Journal of Medicine*, vol. 351, no. 27, pp. 2817-2826, 2004.

[10] L. J. Van 't Veer *et al.*, "Gene expression profiling predicts clinical outcome of breast cancer," *Nature*, vol. 415, no. 6871, pp. 530–536, 2002.

[11] R. Simon, M. D. Radmacher, K. Dobbin, and L. M. McShane, "Pitfalls in the use of DNA microarray data for diagnostic and prognostic classification," *Journal of the National Cancer Institute*, vol. 99, no. 14, pp. 1072–1081, 2009.

[12] N. Mantel, "Evaluation of survival data and two new rank order statistics arising in its consideration," *Cancer Chemotherapy Reports*, vol. 50, no. 3, pp. 163–170, 1966.

[13] R. Peto and J. Peto, "Asymptotically efficient rank invariant test procedures," *Journal of the Royal Statistical Society, Series A*, vol. 135, no. 2, pp. 185–198, 1972.

[14] D. R. Cox, "Regression models and life-tables," *Journal of the Royal Statistical Society: Series B (Methodological)*, vol. 34, no. 2, pp. 187–220, 1972. https://doi.org/10.1111/j.2517-6161.1972.tb00899.x