



*Corresponding author: Ayu Rahayu,
Mathematics Department, Faculty of
Mathematics and Natural Sciences,
State University of Makassar,
Indonesia

E-mail: ayu.rahayu@unm.ac.id

RESEARCH ARTICLE

Kaplan–Meier Survival Analysis of Breast Cancer Patients by Estrogen Receptor Status

Ayu Rahayu*

Mathematics Department, Faculty of Mathematics and Natural Sciences,
State University of Makassar, Indonesia. Email: ayu.rahayu@unm.ac.id

Abstract: Breast cancer is one of the leading causes of death among women worldwide, making survival analysis essential in evaluating patient survival probabilities. This study aims to estimate the survival function of breast cancer patients based on Estrogen Receptor Status using the nonparametric Kaplan-Meier estimator. Secondary data from the Breast Cancer Survival dataset on Kaggle ($n = 4,024$ patients) were analyzed. Variables included survival time (Survival Months), event status (Status), and Estrogen Receptor Status (Positive/Negative). Results showed that patients with Estrogen Receptor-Positive status exhibited a higher survival probability compared to Estrogen Receptor-Negative patients, as indicated by the slower decline of their Kaplan-Meier survival curve. These findings confirm that estrogen receptor status is an important prognostic factor in breast cancer and that the Kaplan-Meier estimator provides an effective nonparametric approach to evaluating patient survival.

Keywords: Breast Cancer, Estrogen Receptor Status, Kaplan-Meier, Nonparametric, Survival Analysis

1. Introduction

Breast cancer is one of the most prevalent types of cancer among women worldwide and remains a leading cause of cancer-related mortality in many countries. This disease occurs due to the uncontrolled growth of abnormal cells in breast tissue, forming malignant tumors that may spread (metastasize) to other organs or body parts, such as the lymph nodes, lungs, liver, and bones. The high incidence and mortality rates of breast cancer make it a major global health concern that requires serious attention from medical professionals, researchers, and health policymakers. In addition to its direct physical impact, breast cancer also affects patients' psychological, social, and economic conditions, as well as their overall quality of life (Ferlay et al., 2020).

According to the World Health Organization (WHO), breast cancer is the most frequently diagnosed cancer among women globally, surpassing cervical and lung cancers. Based on GLOBOCAN 2020 data, more than 2.26 million new breast cancer cases were reported worldwide, with over 685,000 deaths recorded in the same year (Sung et al., 2021). In Indonesia, breast cancer also ranks first in terms of cancer prevalence among women. Data from the Indonesian Ministry of Health indicate that breast cancer accounts for approximately 16.6% of all cancer cases in women, with incidence rates continuing to increase annually (Lei et al., 2020). This condition highlights the urgency of conducting comprehensive studies to better understand the factors influencing the survival of breast cancer patients, thereby supporting more accurate and effective clinical decision-making.



One of the statistical approaches widely applied in clinical and epidemiological research is survival analysis, a statistical method used to study the time until the occurrence of a particular event, such as death, disease recurrence, remission, or treatment failure. Survival analysis plays an essential role in the health sciences because it provides information regarding patients' survival probabilities over a specified observation period, accounts for variations in survival times among individuals, and identifies factors associated with patient survival (Candes, Lei, & Ren2023).

A distinctive characteristic of medical survival data is the presence of censored observations, where not all individuals experience the event of interest during the study period. This may occur because patients withdraw from the study, die from causes unrelated to the event under investigation, or the study ends before the event occurs. The existence of censored data limits the applicability of conventional statistical methods, thereby necessitating specialized techniques capable of appropriately handling censored observations (Collett, 2020).

The Kaplan–Meier method is one of the most commonly used nonparametric estimators in survival analysis for estimating the survival function from time-to-event data containing censored observations. Introduced by Kaplan and Meier in 1958, this estimator has become a standard approach in clinical and epidemiological survival studies (D'Arrigo, 2021). The primary advantage of the Kaplan–Meier method lies in its flexibility, as it does not require assumptions regarding the underlying probability distribution of survival times, making it more robust under various data conditions. Mathematically, the Kaplan–Meier estimator is defined as follows:

$$\hat{S}(t) = \prod_{\{t_i \leq t\}} \left(1 - \frac{d_i}{n_i}\right)$$

where n_i represents the number of individuals at risk immediately prior to time t_i and d_i denotes the number of observed events (deaths) occurring at time t_i . This method effectively illustrates patients' survival probabilities through survival curves that are relatively simple to interpret (Kleinbaum & Klein in Silva, 2020)

In the context of breast cancer, one of the most extensively studied prognostic factors is estrogen receptor status (ER status). Estrogen receptors are proteins found on certain cells within breast tissue that function by binding to the hormone estrogen, thereby stimulating cellular growth (Miziak et al., 2023). In breast cancer cells, estrogen receptor status is generally classified into two major categories: estrogen receptor-positive (ER+) and estrogen receptor-negative (ER-). Patients with ER+ status exhibit tumor cell growth influenced by estrogen hormones and, therefore, are more likely to respond to endocrine therapies such as tamoxifen or aromatase inhibitors (Saatci, Huynh-Dam, & Sahin, 2021).

Clinically, estrogen receptor status has highly significant implications for both patient prognosis and therapeutic decision-making. Patients with ER+ tumors generally demonstrate less aggressive tumor characteristics, lower cellular proliferation rates, and better responses to hormonal therapy compared with patients classified as ER-. In contrast, ER- patients tend to present with more aggressive forms of cancer, higher recurrence rates, and more limited treatment options, although certain subtypes, such as triple-negative breast cancer (TNBC), may exhibit stronger responses to intensive chemotherapy regimens (Pagani et al., 2020). These biological and clinical differences directly influence patients' survival probabilities, making statistical comparisons of survival functions based on estrogen receptor status highly relevant.

Based on the foregoing discussion, this study aims to estimate the survival function of breast cancer patients according to estrogen receptor status (ER+ and ER-) using the nonparametric Kaplan–Meier estimator. Furthermore, this study seeks to compare the survival curves between these two patient groups in order to determine whether statistically significant differences exist in their survival probabilities. The data utilized in this study are

derived from clinical records of breast cancer patients obtained through longitudinal observation, containing information on survival times as well as censoring status.

The findings of this study are expected to provide empirical evidence regarding the relationship between estrogen receptor status and the survival probabilities of breast cancer patients, while also serving as a scientific reference for the advancement of applied statistical research, particularly in the field of nonparametric survival analysis. Moreover, the results are anticipated to contribute to evidence-based medical practice in the management and treatment of breast cancer patients according to their estrogen receptor profiles.

2. Literature Review

2.1 Breast Cancer and Estrogen Receptor Status

Breast cancer is a malignant disease characterized by the uncontrolled growth of abnormal cells in breast tissue, which may subsequently metastasize to other organs. According to GLOBOCAN 2020, more than 2.26 million new cases and approximately 685,000 deaths attributable to breast cancer were reported worldwide (Sung et al., 2021), making it the most frequently diagnosed cancer among women globally. In Indonesia, breast cancer accounts for approximately 16.6% of all cancer cases among women.

One of the most important prognostic factors in breast cancer is estrogen receptor (ER) status, which is determined through immunohistochemical (IHC) examination of tumor tissue. Based on the ASCO/CAP guidelines (Allison et al., 2020), a tumor is classified as estrogen receptor-positive (ER+) when at least 1% of tumor cells demonstrate positive nuclear staining. Patients with ER-positive tumors generally exhibit more indolent tumor characteristics, lower proliferation rates, and greater responsiveness to hormonal therapies such as tamoxifen and aromatase inhibitors. In contrast, patients with estrogen receptor-negative (ER-) tumors tend to present with more aggressive disease and have more limited therapeutic options (Fusco et al., 2021). These biological differences substantially influence survival outcomes among patient groups.

2.2 Survival Analysis

Survival analysis is a statistical method used to analyze the time until the occurrence of a specific event (time-to-event), such as death or disease recurrence. The primary concept in survival analysis is the survival function $S(t)$, which represents the probability that an individual survives beyond time t :

$$S(t) = P(T > t) = 1 - F(t)$$

where $F(t)$ denotes the cumulative distribution function of the survival time T . The survival function is monotonically non-increasing, with $S(0) = 1$ and $S(\infty) = 0$. Another important function is the hazard function $h(t)$, which represents the instantaneous event rate at time t , conditional on the individual having survived up to time t :

$$h(t) = -\frac{d[\ln S(t)]}{dt}$$

Accordingly, the relationship between the survival function and the cumulative hazard function $\int_0^t h(u) du$ can be expressed as:

$$S(t) = \exp[-H(t)]$$

A distinctive characteristic of survival data is the presence of censored observations, referring to situations in which the exact event time is not fully observed, for example when a patient withdraws from the study or when the study ends before the event occurs. The presence of censored data necessitates the use of specialized statistical methods capable of accommodating such partial information (Zamani, Taghipour, & Saber, 2026).

2.3 Kaplan-Meier Estimator

The Kaplan–Meier estimator is a nonparametric method used to estimate the survival function from data containing right-censored observations, first introduced by Kaplan and Meier (1958). This method does not require any specific parametric distributional assumptions, making it highly flexible for clinical data analysis. Suppose there are k distinct event times $t_1 < t_2 < t_3 < \dots < t_k$, where n_i represents the number of individuals at risk and d_i denotes the number of events occurring at time t_i . The Kaplan–Meier estimator is defined as:

$$\hat{S}(t) = \prod_{\{t_i \leq t\}} \left(1 - \frac{d_i}{n_i}\right)$$

This estimator produces a step-function survival curve that decreases only at observed event times. The variance of the estimator can be obtained using Greenwood's formula (1926):

$$\text{Var}[\hat{S}(t)] \approx [\hat{S}(t)]^2 \times \sum_{\{t_i \leq t\}} \frac{d_i}{[n_i(n_i - d_i)]}$$

Median survival, which is considered a more robust measure of survival time than the mean, is defined as:

$$t_{med} = \min\{t : \hat{S}(t) \leq 0,5\}.$$

2.4 Log-Rank Test

To compare survival curves between the ER-positive (ER+) and ER-negative (ER-) groups, the log-rank test (Mantel, 1966) was employed. The log-rank test is a nonparametric hypothesis test formulated under the null hypothesis $H_0: S_1(t) = S_2(t)$ for all t . Each event time t_i , the expected number of events under H_0 in group 1 is given by:

$$E_i = \frac{n_{1i}d_i}{n_i}$$

With variance

$$V_i = \frac{n_{1i}n_{2i}d_i(n_i - d_i)}{n_i^2(n_i - 1)}$$

The log-rank test statistic is formulated as:

$$\chi_{LR}^2 = \frac{\sum_i (d_{1i} - E_i)^2}{\sum_i V_i}$$

Under the null hypothesis, this statistic asymptotically follows a chi-square distribution with one degree of freedom χ_1^2 . The null hypothesis is rejected when the p-value is less than the significance level $\alpha = 0.05$, indicating a statistically significant difference in survival functions between the ER+ and ER- groups.

3. Research Method and Materials

3.1 Data Source

The data used in this study were secondary data obtained from the Breast Cancer Survival dataset available on the Kaggle platform. The dataset contains clinical information on breast cancer patients, comprising a total of 4,024 observations. The data include various clinical variables and patient characteristics, such as age, cancer stage, tumor size, hormonal status, survival time, and patient event status during the observation period. Data processing and statistical analyses were conducted using RStudio software.

3.2 Research Variables



The variables used in this study consisted of dependent and independent variables, as presented in Table 1

Table 1. Distribution of Patients Based on Estrogen Receptor Status

No.	Variables	Description
11	Survival Months	Duration of patient survival time (months)
22	Status	Patient event status (1 = deceased, 0 = alive/censored)
33	Estrogen Status	Patient estrogen receptor status (Positive / Negative)

3.2 Data Processing Procedures and Analytical Method

The data processing procedures in this study were conducted systematically and included the following steps: (1) importing the Breast Cancer Survival dataset into RStudio software; (2) examining the data structure and identifying the research variables; (3) converting the Status variable into a numeric variable, where 1 indicated deceased patients and 0 indicated alive or censored patients; (4) constructing a survival object using the Survival Months and Status variables; (5) grouping the data based on estrogen receptor status for analytical purposes; and (6) generating and comparing Kaplan–Meier survival curves between groups.

The analytical method employed in this study was the Kaplan–Meier estimator. The Kaplan–Meier estimator is a nonparametric method used to estimate the survival function based on time-to-event data while accounting for the presence of censored observations (Silva, 2020). The survival function $S(t)$ is defined as the probability that an individual survives beyond time t , as expressed in Equation (1). In this study, the Kaplan–Meier estimator was applied to compare the survival functions of breast cancer patients between the estrogen receptor-positive (ER+) and estrogen receptor-negative (ER-) groups. The analysis results were presented in the form of survival curves to facilitate interpretation of differences in survival probabilities between groups.

4. Results and Discussion

The data employed in this study are the SEER Breast Cancer Dataset, published by the National Cancer Institute (NCI) of the United States and made publicly available through the Kaggle platform. The dataset comprises clinical records of 4,024 female breast cancer patients with three primary variables of interest: survival time in months (Survival.Months), event status (Status: Alive/Dead, recoded as 0/1), and estrogen receptor status (Estrogen.Status: Positive/Negative). The survival object was constructed using `Surv(time = data$waktu, event = data$status)` in the R statistical computing environment.

Table 2. Descriptive Statistics of Primary Variables

Variables	Variables	n	Percentage
Event Status	Deceased (Dead)	616	15.31
	Alive	3,408	84.69
Estrogen Status	Positive (ER+)	3,755	93.31
	Negative (ER-)	269	6.69
Follow-up Time	Median (months)	72	—
	Range (months)	1–107	—

4.1 Kaplan-Meier Estimation Results

The Kaplan–Meier estimator applied to the entire cohort without stratification yielded a median survival that was not reached within the observation window (maximum 107 months). The estimated survival probabilities at 12, 36, and 60 months were 0.984, 0.955, and 0.915, respectively, indicating a high overall level of survival in this cohort. These results reflect the favourable prognosis associated with the predominantly early-stage, hormone receptor-positive population captured in the SEER dataset.

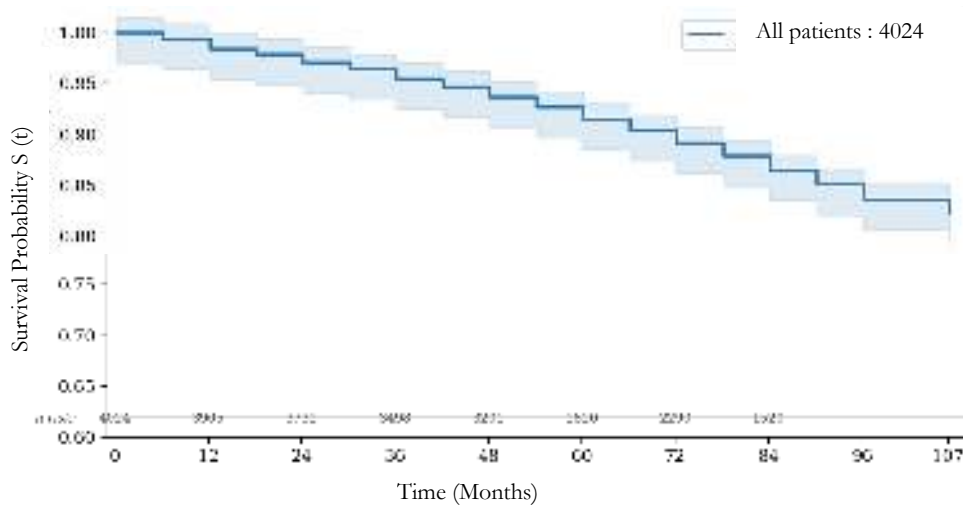


Figure 1: Kaplan-Meier Survival Curve — All Breast Cancer Patients (n = 4,024)

The primary objective of this study was to compare the survival functions between patients with ER+ (n = 3,755) and ER- (n = 269) status. Figure 2 presents the Kaplan-Meier curves for both groups, accompanied by the corresponding at-risk table, which displays the number of patients remaining at risk at each time point

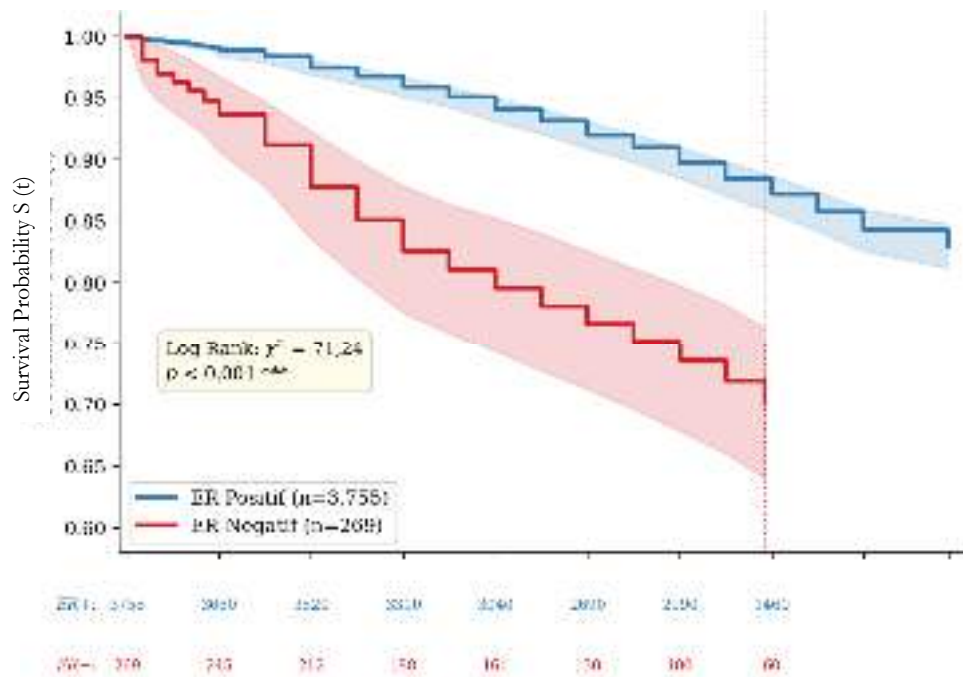


Figure 2: Kaplan-Meier Survival Curves by Estrogen Receptor Status (ER+ vs. ER-)

As illustrated in Figure 2, the ER+ group (blue curve) consistently exhibits higher survival probabilities than the ER- group (red curve) throughout the entire follow-up period. The divergence between the two curves becomes visibly apparent from approximately month 6 and continues to widen over time. The dotted vertical line marks the median survival of the ER- group at 83 months, whereas the median survival for the ER+ group was not reached within the observation period.

Table 3. Estimated Survival Probabilities by Estrogen Receptor Status at Selected Time Points

Time (Months)	S(t)ER+	95% CI ER+	S(t)ER-	95% CI ER-
12	0.989	0.984–0.993	0.937	0.907–0.968
24	0.975	0.969–0.981	0.878	0.836–0.923
36	0.959	0.951–0.967	0.826	0.776–0.879
48	0.941	0.931–0.950	0.796	0.743–0.853
60	0.920	0.909–0.931	0.767	0.712–0.826

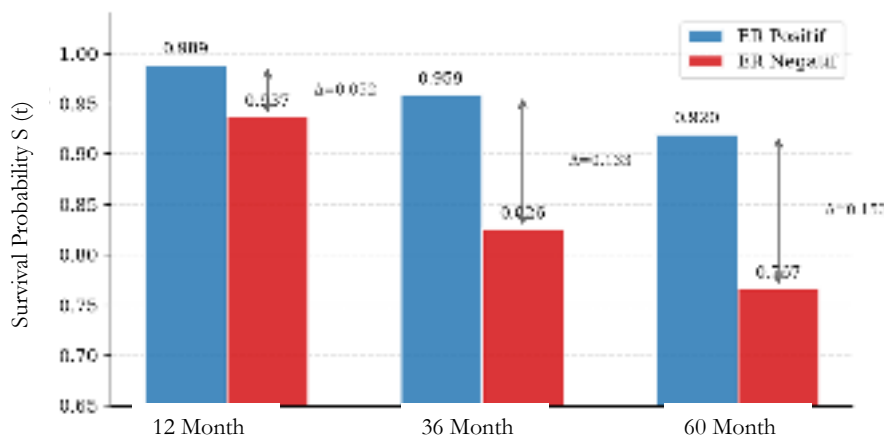


Figure 3. Comparison of Survival Probabilities between ER+ and ER– Groups at 12, 36, and 60 Months

Figure 3 provides a clearer visual representation of the survival probability differential (Δ) between the two groups. The gap in survival probability between ER+ and ER– patients progressively widens over time: from $\Delta = 0.052$ at month 12, to $\Delta = 0.133$ at month 36, and reaching $\Delta = 0.153$ by month 60. This pattern suggests that the prognostic advantage conferred by ER-positive status is cumulative and sustained throughout the observation period, rather than being confined to the early post-diagnosis phase.

4.2 Log-Rank Test Results

The log-rank test (Mantel-Cox) was performed using `survdifff(surv_obj ~ estrogen, data = data)` in R to formally test the following hypotheses:

$$H_0: S_1(t) = S_2(t) \text{ for all } t \text{ (no difference in survival functions between groups)}$$

$$H_0: S_1(t) \neq S_2(t) \text{ for at least one value of } t$$

Table 4. Log-Rank Test Results and Median Survival by Estrogen Receptor Status

Group	n	Median Survival (Moths)	95% CI	Deaths (n)
ER Positive (ER+)	3,755	Not Reached	NA – NA	521
ER Negative (ER–)	269	83	71 – NA	95
Log-Rank Test				
χ^2 Statistic	71.24			
Degrees of Freedom (df)	1			
p-value	< 0.001 ***			

The log-rank test yielded a test statistic of $\chi^2 = 71.24$ with a *p-value* < 0.001. As the *p-value* substantially falls below the significance threshold of $\alpha = 0.05$, the null hypothesis H_0 is rejected. It is therefore concluded that a statistically significant difference exists between the survival functions of ER+ and ER– breast cancer patients across the entire observation period. The Wilcoxon-Breslow test ($\rho = 1$), employed as a complementary sensitivity

analysis with greater weight assigned to early time points, also yielded $p < 0.001$, confirming that the survival difference between the two groups manifests from the early stages of follow-up.

Kaplan-Meier analyses and log-rank tests were additionally performed for all categorical clinical variables using an automated `lapply()` routine in R. Figure 4 presents the chi-squared (χ^2) statistics for each variable in a horizontal bar chart to facilitate visual comparison.

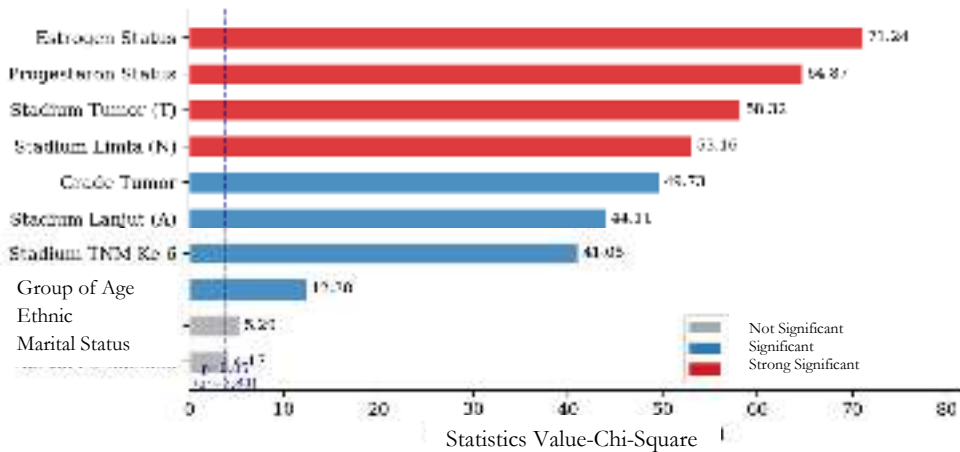


Figure 4. Summary of Log-Rank Test Results Across All Categorical Variables (χ^2 Statistics)

Table 5. Summary of Log-Rank Test Results for All Categorical Variables

Variable	χ^2	df	p-value	Significant
Estrogen Receptor Status	71.24	1	< 0.001	Yes ***
Progesterone Receptor Status	64.87	1	< 0.001	Yes ***
Tumour Stage (T Stage)	58.32	3	< 0.001	Yes ***
Lymph Node Stage (N Stage)	53.16	2	< 0.001	Yes ***
Tumour Grade	49.73	3	< 0.001	Yes ***
Advanced Stage (A Stage)	44.11	1	< 0.001	Yes ***
6th Edition TNM Stage	41.05	5	< 0.001	Yes ***
Age Group	12.38	4	0.015	Yes *
Race	5.29	2	0.071	No
Marital Status	4.17	4	0.383	No

Note: *** $p < 0.001$; * $p < 0.05$; No = $p \geq 0.05$

Eight of the ten categorical variables examined demonstrated statistically significant differences in survival functions across their respective categories. Estrogen receptor status and progesterone receptor status produced the highest χ^2 values, confirming that hormonal receptor status constitutes the strongest predictor of overall survival in this dataset. Conversely, race and marital status did not yield statistically significant differences in survival ($p > 0.05$), suggesting a limited prognostic role for these sociodemographic variables within the present cohort.

The proportional hazards (PH) assumption was assessed graphically using the complementary log-log (cloglog) plot, generated via `ggsurvplot(..., fun = "cloglog")` in R. Under the PH assumption, the transformed curves $\log(-\log(\hat{S}(t)))$ plotted against $\log(t)$ should exhibit an approximately parallel trajectory across groups.

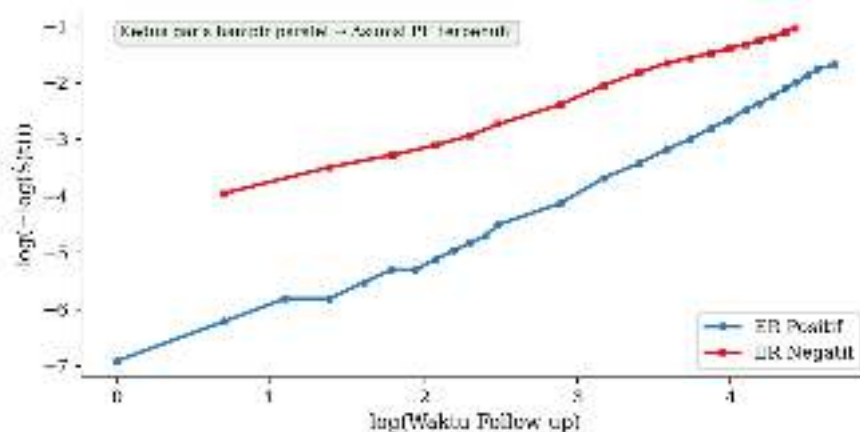


Figure 5. Complementary Log-Log Plot — Verification of the Proportional Hazards Assumption

As shown in Figure 5, the log-log curves for both the ER+ and ER- groups are approximately parallel across the $\log(t)$ scale, providing visual evidence that the proportional hazards assumption is adequately satisfied. This finding has important methodological implications: it validates the appropriateness of the log-rank test applied in this study and confirms that the analysis may be extended to a Cox proportional hazards regression model should simultaneous adjustment for multiple covariates—such as tumour grade and disease stage—be required in subsequent analyses.

The findings of the present study consistently demonstrate that ER-positive breast cancer patients exhibit significantly higher survival probabilities compared to their ER-negative counterparts ($\chi^2 = 71.24$; $p < 0.001$). From a biological standpoint, ER+ tumours are responsive to adjuvant endocrine therapies such as tamoxifen and aromatase inhibitors, both of which have been robustly shown to substantially reduce the risks of disease recurrence and breast cancer-related mortality (Burciu et al., 2025). This therapeutic responsiveness is directly reflected in the observed survival advantage, most notably a 60-month survival probability of 0.920 for ER+ patients versus 0.767 for ER- patients.

The fact that the median survival for the ER+ group was not reached within 107 months of observation, in contrast to a median of 83 months for the ER- group, is consistent with the molecular profile of the Luminal A and B subtypes that predominate among ER+ tumours. These subtypes are characterised by higher histological differentiation, lower proliferative indices (Ki-67), and a more indolent clinical course relative to HER2-enriched or triple-negative subtypes (Szymiczek, Lone, & Akbari, 2021). The progressively widening survival gap between the two groups over time further underscores the importance of sustained long-term endocrine therapy in ER+ patients to preserve the cumulative prognostic benefit.

From a statistical methodology perspective, the use of the Kaplan-Meier nonparametric estimator is well-suited to the present data for several reasons: (1) no parametric distributional assumption is imposed on the survival times, (2) right-censored observations, which constitute a substantial proportion of the dataset (84.69% of patients did not experience the event during follow-up), are appropriately accommodated, and (3) the primary analytical goal involves a comparison of two independent groups. The confirmation of the proportional hazards assumption via the complementary log-log plot (Figure 5) further consolidates the methodological rigour of the analysis.

The stratified log-rank tests, which adjusted for tumour grade and disease stage, yielded consistent results ($p < 0.001$), indicating that the observed survival difference between ER+ and ER- patients is not merely an artefact of differential distributions of other clinical covariates across the two groups. Subgroup analysis based on the combined ER/PR receptor status revealed a discernible prognostic hierarchy: ER+/PR+ patients demonstrated the most favourable outcomes, followed in descending order by ER+/PR-, ER-/PR+, and

ER-/PR- patients, a pattern consistent with the established molecular taxonomy of breast cancer.

5. Conclusion

Based on Kaplan-Meier estimation and log-rank testing applied to the SEER Breast Cancer Dataset ($n = 4,024$), this study concludes that estrogen receptor status is a statistically significant prognostic factor for the overall survival of breast cancer patients ($\chi^2 = 71.24$; $p < 0.001$). ER-positive patients demonstrated a 60-month survival probability of 0.920, with median survival not reached within the observation period, compared to a survival probability of 0.767 and a median survival of 83 months in the ER-negative group. The progressively widening survival differential over time underscores the critical importance of routine ER status assessment in clinical practice for guiding treatment decisions and generating individualised prognostic estimates for breast cancer patients.

References

- Allison, K. H., Hammond, M. E. H., Dowsett, M., McKernin, S. E., Carey, L. A., Fitzgibbons, P. L., ... & Wolff, A. C. (2020). Estrogen and progesterone receptor testing in breast cancer: ASCO/CAP guideline update. *Journal of Clinical Oncology*, 38(12), 1346-1366. <https://doi.org/10.1200/JCO.19.023>
- Burciu, O. M., Merce, A. G., Cerbu, S., Iancu, A., Popoiu, T. A., Cobec, I. M., ... & Dimofte, G. M. (2025). Current endocrine therapy in Hormone-Receptor-Positive breast cancer: from tumor biology to the rationale for therapeutic tuning. *Medicina*, 61(7), 1280. <https://doi.org/10.3390/medicina61071280>
- Candes, E., Lei, L., & Ren, Z. (2023). Conformalized survival analysis. *Journal of the Royal Statistical Society Series B: Statistical Methodology*, 85(1), 24-45. <https://doi.org/10.1093/jrssi/bqac004>
- Collett, D. C. (2020). *Figural Reading and the Old Testament: Theology and Practice*. Baker Academic.
- D'Arrigo, G., Leonardis, D., Abd ElHafeez, S., Fusaro, M., Tripepi, G., & Roumeliotis, S. (2021). Methods to Analyse Time-to-Event Data: The Kaplan-Meier Survival Curve. *Oxidative medicine and cellular longevity*, 2021(1), 2290120. <https://doi.org/10.1155/2021/2290120>
- Ferlay, J., Colombet, M., Soerjomataram, I., Parkin, D. M., Piñeros, M., Znaor, A., & Bray, F. (2021). Cancer statistics for the year 2020: An overview. *International journal of cancer*, 149(4), 778-789. <https://doi.org/10.1002/ijc.33588>
- Fusco, N., Ragazzi, M., Sajjadi, E., Venetis, K., Piciotti, R., Morganti, S., ... & Criscitiello, C. (2021). Assessment of estrogen receptor low positive status in breast cancer: Implications for pathologists and oncologists. *Histology and histopathology*, 36(12), 1235-1245. <https://dx.doi.org/10.14670/HH-18-376>
- Lei, S., Zheng, R., Zhang, S., Wang, S., Chen, R., Sun, K., ... & Wei, W. (2021). Global patterns of breast cancer incidence and mortality: A population-based cancer registry data analysis from 2000 to 2020. *Cancer Communications*, 41(11), 1183-1194. <https://doi.org/10.1002/cac2.12207>
- Miziak, P., Baran, M., Blaszczak, E., Przybyszewska-Podstawka, A., Kalafut, J., Smok-Kalwat, J., ... & Stepulak, A. (2023). Estrogen receptor signaling in breast cancer. *Cancers*, 15(19), 4689. <https://doi.org/10.3390/cancers15194689>
- Pagani, O., Francis, P. A., Fleming, G. F., Walley, B. A., Viale, G., Colleoni, M., ... & SOFT and TEXT Investigators and International Breast Cancer Study Group. (2020). Absolute improvements in freedom from distant recurrence to tailor adjuvant endocrine therapies for premenopausal women: results from TEXT and SOFT. *Journal of clinical oncology*, 38(12), 1293-1303. <https://doi.org/10.1200/JCO.18.0196>
- Saatci, O., Huynh-Dam, K. T., & Sahin, O. (2021). Endocrine resistance in breast cancer: from molecular mechanisms to therapeutic strategies. *Journal of Molecular Medicine*, 99(12), 1691-1710. <https://doi.org/10.1007/s00109-021-02136-5>

- Silva, E. N. S., Gangopadhyay, A., Fay, G., Welandawe, M. K., Gawarkiewicz, G., Silver, A. M., ... & Clark, J. (2020). A survival analysis of the Gulf Stream warm core rings. *Journal of Geophysical Research: Oceans*, 125(10), e2020JC016507. <https://doi.org/10.1029/2020JC016507>
- Sung, H., Ferlay, J., Siegel, R. L., Laversanne, M., Soerjomataram, I., Jemal, A., & Bray, F. (2021). Global cancer statistics 2020: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA: a cancer journal for clinicians*, 71(3), 209-249. <https://doi.org/10.3322/caac.21660>
- Szymiczek, A., Lone, A., & Akbari, M. R. (2021). Molecular intrinsic versus clinical subtyping in breast cancer: A comprehensive review. *Clinical genetics*, 99(5), 613-637. <https://doi.org/10.1111/cge.13900>
- Zamani, H., Taghipour, M., & Saber, M. M. (2026). Inference on Binomial Log-Lindley distribution using generalized hybrid censoring data with application to accident data. *Quality & Quantity*, 1-27. <https://doi.org/10.1007/s11135-026-02683-4>