https://doi.org/10.31272/jae.i149.1395
https://admics.uomustansiriyah.edu.iq/index.php/admeco
P-ISSN: 1813–6729 E-ISSN: 2707-1359

Restricted Mean Survival Time Approaches Implications for Breast Cancer Data

Sami A. Obed
Dept of Statistics & Informatics, College of Administration and Economics, Salahaddin University, Erbil, Kurdistan, Iraq
Email: sami.obed@su.edu.krd, ORCID ID:/ https://orcid.org/0000-0002-2866-5886
Kurdistan I. Mawlood
Dept of Statistics & Informatics, College of Administration and Economics, Salahaddin University, Erbil, Kurdistan, Iraq
Email: kurdistan.mawlood@su.edu.krd , ORCID ID :/ https://orcid.org/0000-0002-1612-1996

Article Information

Article History:
Received: 16 / 03 / 2025
Accepted: 24 / 07 / 2025
Available Online: 01 / 09 / 2025
Page no: 30 – 41

Keywords:
RMST, Breast Cancer, Survival Analysis, Survival Curve, CIF.

Correspondence:
Researchers name:
Kurdistan Ibrahim Mawlood

Email: kurdistan.mawlood@su.edu.krd

Abstract

This research assesses the utility of Restricted Mean Survival Time (RMST) as a viable alternative to conventional survival analysis techniques, including the Kaplan-Meier test and Cox Proportional Hazards model, especially when the proportional hazards assumption is not met. RMST offers an alternative approach to using hazard ratios in such situations. This study aimed to evaluate the impact of predictive factors on breast cancer survival. The Restricted Mean Survival Time (RMST) offers a clinically interpretable measure of average survival time within a defined period, providing a more detailed understanding of treatment effects and patient outcomes. Additionally, the study utilised the Cumulative Incidence Function (CIF) model and probability estimates for each time point to further analyse survival data.

Using data from 4,420 breast cancer patients at Rizgary Hospital, this research quantifies the impact of various prognostic factors on survival outcomes. The study demonstrates that RMST effectively captures differences in survival probabilities across treatment groups, with patients undergoing surgery or receiving radiotherapy showing significantly higher survival rates than those who did not. Additionally, the analysis reveals that factors such as family history and tumour grade significantly influence survival outcomes, highlighting the heterogeneity of breast cancer. Statistical comparisons using RMST, alongside traditional tests like the Log-rank and Gehan-Wilcoxon tests, confirm significant differences in survival curves, particularly in early-stage survival, emphasising the importance of timely interventions.

1. Introduction:

Survival times represent time-to-event outcomes that analyse the occurrence of specific events within defined temporal parameters. Consequently, the restricted mean survival time (RMST) methodology not only yields survival duration but also estimates the average life expectancy or treatment duration for specific patient subgroups.[14]

Breast cancer is one of the most common types of cancer affecting women worldwide, though it can also occur in men. The disease occurs when breast cells grow uncontrollably, forming a tumour that can often be detected through imaging or physical examination. The Restricted Mean Survival Time (RMST) approach has gained attention for its ability to provide meaningful estimates of patient survival while accounting for censored data.[12][13]

RMST calculates the average time until an event occurs, for a predefined time horizon, and can be

particularly useful when analysing therapies that can prolong survival without necessarily altering the risk of death.

In clinical practice, the RMST approach allows health professionals to deliver personalised prognostic information adapted to the circumstances of individual patients. By presenting RMST estimates, doctors can effectively communicate the duration of the expected average survival for patients, given their specific treatment plans and cancer characteristics. For example, a patient with breast cancer who considers chemotherapy can find it more satisfactory to understand the average life expectancy, factored by possible delays related to the treatment in the progression of the disease.[2]

2. Objective of the Study: -

This study aimed to assess the applicability and effectiveness of Restricted Mean Survival Time (RMST) in breast cancer patients. RMST offers a straightforward summary of survival time across different groups (e.g., treatment vs. control) by comparing average survival times within a specified period. This makes it more intuitive to interpret than hazard ratios, which focus on relative risks. Estimate the probability of specific events occurring over time in the presence of competing risks. Identify which factors significantly influence outcomes in breast cancer patients (e.g., age, gender, tumour grade, family history, …) by the cumulative incidence function (CIF) model of the event of interest.

3. Overview of Restricted Mean Survival Time (RMST) and Its Relevance in Breast Cancer Research: -

In breast cancer research, the Restricted Mean Survival Time (RMST) is a vital metric for evaluating patient outcomes, offering a nuanced alternative to traditional survival analysis methods, unlike standard measures that often emphasise survival rates. RMST quantifies the expected survival time within a specified period, accounting for censoring and treatment effects. This is especially important in healthcare settings where survival and quality of life are closely related. Understanding the advantages of treatment can be improved by including RMST in decision-making, particularly when longer progression-free survival is not always associated with improved survival rates. [3][14]

4. Survival analysis:

Time-to-event analysis focuses on measuring the duration until a specific event occurs. Examples include the time until infection, disease recurrence, or recovery in health sciences; Survival analysis, which is a key component of reliability studies in engineering, often refers to these durations as "failure times," as the primary focus is on how long a component operates correctly before it fails.[1][7]

Survival analysis encompasses parametric, semiparametric, and nonparametric approaches. Certain statistical probability distributions are well-suited for modelling survival times. Among the most frequently used distributions are the exponential, Weibull, and lognormal.[12][2]

5. Survival Function

The survival function represents the likelihood of surviving beyond a specific point in time. It indicates the probability that an individual's survival time will exceed a particular value. Since the cumulative distribution function, F(t), measures the probability that the survival time is less than or equal to a specific point of time, the survival function for a continuous distribution, S(t), is defined as the complement of the cumulative distribution function as in equation (1):

$$S(t) = P(T > t) \quad (1)$$

Where:

T: is the random variable representing the time to the event of interest (e.g, death, failure, recovery).
t: is a specific time point.[11]

5.1 Properties of the Survival Function:

1. **Starts at 1**: At time t=0, the survival probability is one because no events have occurred yet.
S(0)=1

2. **Non-Increasing**: The survival function decreases (or stays constant) over time as more events occur, as defined in equation (2):

$$S(t_2) \leq S(t_1) \quad \text{for } t_2 > t_1 \tag{2}$$

3. **Approaches 0**: As t→∞, S(t) approaches 0 (assuming all individuals will eventually experience the event). [6]

6. Applications of the Survival Function: -

1. **Estimating survival probabilities**: How likely is it for a patient or system to survive beyond a specific time?
2. **Comparing groups**: Assessing the difference in survival probabilities between treatment groups, populations, or conditions.
3. **Modelling risk**: Supporting the evaluation of interventions or risk factors.[10]

7. Censoring:

Censoring is a critical concept in survival analysis (e.g., time to death, failure, recovery). Censoring occurs when the complete information about the survival time of an individual is not available. This can happen for various reasons, such as a study ending before the event occurs, loss to follow-up, or withdrawal from the study. Censoring has three types (Right, Left and interval).[4]

8. Restricted Mean Survival Time:

Restricted Mean Survival Time (RMST) is a key metric in survival analysis, often employed to evaluate and compare the effectiveness of treatments or interventions in clinical trials. It represents the expected time a subject will survive within a specified time interval, often called the "restriction time" or "truncation time." Unlike measures like median survival time or hazard ratios, RMST provides a more interpretable and comprehensive summary of the survival experience within a limited time horizon.[8][10]

Let *T* be a nonnegative random variable that represents the failure time of an individual from a homogeneous population. The survivor function (also known as the survival function) of *T* is defined as equation (1).

Assume that τ is a prespecified time point of interest. Let R be the minimum of T and τ:

R = min. (T;)

The restricted mean survival time is defined as the expected value of *R*:

RMST (τ)= E (R)= E [min (T; τ)]

$$RMST(\tau) = \int_0^{\tau} S(t)dt \tag{3}$$

The area under the survivor function can be evaluated it (4) over [0, τ]

Restricted mean survival time (RMST) holds significant promise as a valuable and sensitive measure for analysing trial data involving time-to-event outcomes.[5][9]

8.1 Mean Survival Time Compared to Restricted Mean Survival Time:

For every study subject, the traditional mean survival time is the average time that passes until an event occurs. To overcome this problem, RMST limits the study to a particular time horizon, which results in a more comprehensive summary of survival up to that point.[1]

The restricted mean survival time, denoted as $\mu(\tau)$, for a random variable T, represents the average of the minimum values between (T, t^*). It can be calculated as the area under the survival curve S(t) from time zero up to τ Is defined as a function (4):

$$Mu\mu(\tau) = E(\min(T,\tau)) = \int_0^{\tau} S(t)dt \tag{4}$$

Where $S(t)$ This is the all-cause survival function. T is time to death. $\mu(\tau)$as the τ year life expectancy.[12][13]

8.2 RMST has attractive properties:

1- It does not depend on the proportional hazard's assumption;

2- It can effectively summarize differences in survival, even when survival curves initially separate.
3- Unlike hazard ratios, which focus on relative risks, RMST provides insights into absolute risk levels.[1]

9. RMST Estimate in Weibull Distribution as equation (5):

$$\text{RMST} = \int_0^{\tau} \exp\{-(\alpha t^{\beta})\}\ dt \tag{5}$$

$$\text{Let y} = (\alpha t^{\beta}) \rightarrow\ \text{t}^{\beta} = \frac{y}{\alpha}\ ,\ \text{t} = (\frac{y}{\alpha})^{\frac{1}{\beta}}\ \ ,d_t = \frac{1}{\beta}\left(\frac{y}{\alpha}\right)^{\frac{1}{\beta}-1} * \frac{1}{\alpha} dy \tag{6}$$

$$\text{RMST} = \int_0^{\alpha t^{\beta}} e^{-y}\ \ \frac{1}{\beta\alpha}\ \ \frac{y^{\frac{1}{\beta}-1}}{\alpha^{\frac{1}{\beta}-1}}\ dy, \qquad \text{RMST} = \int_0^{\alpha t^{\beta}} e^{-y}\ \ \frac{1}{\beta\alpha^{\frac{1}{\beta}}}\ \ y^{\frac{1}{\beta}-1}\ dy \tag{7}$$

$$\text{RMST} = \frac{1}{\beta\alpha^{\frac{1}{\beta}}} \int_0^{\alpha t^{\beta}} e^{-y}\ y^{\frac{1}{\beta}-1}\ dy \tag{8}$$

Use incomplete Gamma Function:

$$\Gamma(s,x) = \int_0^x t^{s-1}\, e^{-t}\, dt \tag{9}$$

Where:

$$s = \frac{1}{\beta}\text{, and x} = \alpha t^{\beta}, \qquad \int_0^{\alpha t^{\beta}} e^{-y}\ y^{\frac{1}{\beta}-1}\ dy = \Gamma\left(\frac{1}{\beta}, \alpha t^{\beta}\right) \tag{10}$$

$$\text{RMST} = \frac{1}{\beta\alpha^{\frac{1}{\beta}}}\Gamma\left(\frac{1}{\beta}, \alpha t^{\beta}\right) = \frac{1}{\beta\alpha^{\frac{1}{\beta}}} \int_0^{\tau} e^{-\alpha t^{\beta}}\ t^{\frac{1}{\beta}-1}\ dt \tag{11}$$

10. Cumulative Incidence Function:

In the study of survival, a statistical technique called a Cumulative Incidence Function (CIF) is used to summaries the likelihood that an event will occur over time when there are competing risks. An individual is at risk of experiencing multiple types of occurrences, and the occurrence of one event precludes the occurrence of the others. This is known as competing risks.[9]

10.1 Calculation of Cumulative Incidence:

The cumulative incidence at time *t* is calculated using the method (6):

$$(t) = \sum_{t_i<t} \left(\frac{\text{Number of Events at } time\ t_i}{Number\ of\ Events\ at\ Risk\ at\ time\ t_i}\right) * \left(1 - \frac{\text{Number of competing Events at } time\ \ t_i}{Number\ of\ Events\ at\ Risk\ at\ time\ t_i}\right) \tag{12}$$

This formula accounts for the fact that individuals who experience a competing event are no longer at risk for the event of interest.

- The CIF is different from the Kaplan-Meier estimate, which does not account for competing risks and may overestimate the probability of the event of interest.
- The CIF is always non-decreasing and ranges between 0 and 1.[5]

10.2 Cumulative Incidence Model:

A statistical framework for examining the likelihood of an event happening over time, especially when there are unsuitable hazards, is called a cumulative incidence model. It is frequently used in epidemiology, medical research, and survival analysis to calculate the probability of a particular event (such as the development of a disease, death, or failure) while taking into consideration other factors that could make the occurrence of interest unlikely as from equation (7):

$$\boldsymbol{\phi}\left(\boldsymbol{F_1}(\boldsymbol{t}|\boldsymbol{z})\right) = \boldsymbol{\alpha}(\boldsymbol{t}) + \boldsymbol{\beta}(\boldsymbol{t})^{\boldsymbol{T}}\boldsymbol{z} \tag{13}$$

Where: ϕ: is the link function. $\alpha(t)$: the baseline functions. $\beta(t)$: is a vector of time-dependent coefficients, and *z* is a vector of covariates. T: denotes the transpose of the vector.[9][3]

11. Results and Discussion

The data used in this research were obtained from the official database of the Rizgary Hospital, which was gathered for patients via direct communication between the patients and the specialist physician. Data from 4420 Breast cancer patients at Nanakali, Rizgary, and Awat three hospitals in Erbil, Iraq's Kurdistan Region, were gathered for this study. During the five years beginning on January 1, 2019, and continuing until August 31, 2024, for all patients with Breast cancer, the survival time was expressed in months, and the data included the following 14 variables. The data were analysed within a survival analysis framework using R and STATA programs.

Table (1): The explanatory variables measured for these data at diagnosis

Name of variables	Description	Percentage	Name of variables	Description	Percentage
Age	<= 5	0.02%	**Occupation**	Wife house	42.60%
	15-Jun	0.05%		Worker	0.05%
	16 - 25	0.50%		Farmer	0.02%
	26 - 35	8.60%		Employee	8.90%
	36 - 45	26.10%		Craftsman	0.10%
	46 - 55	29.80%		Child	0.10%
	56 - 65	21.00%		Retire	48.30%
	66 - 75	10.40%			
	76 - 85	3.10%			
	86 - 95	0.40%			
Gender	Male	1%	**Chemo**	No	72.90%
	Female	99%		Yes	27.10%
Grade	Grade I	4.90%	**Radio**	No	41.10%
	Grade II	49.70%		Yes	58.90%
	Grade III	29.90%			
	Grade IV	15.50%			
Literately	Right	38.50%	**Hormone**	No	79.50%
	Left	59.90%		Yes	20.50%
	Bilateral	1.30%			
	Not Applicate	0.30%			
Family History	No	55%	**Isotope**	No	61.20%
	Yes	45%		Yes	38.80%
Nationality	Iraqi	96.20%	**Target**	No	54.50%
	Arabi	3.50%		Yes	45.50%
	External	0.30%			
Surgery	No	59.70%	**Status**	Death	10.10%
	Yes	40.30%		Alive	89.90%

The table above provides a detailed description of various explanatory variables measured at the time of diagnosis for a specific dataset. Each variable is categorised, and the percentage of each category within the dataset is provided. The dataset is primarily female, with most patients in the 36-55 age range. The most common tumour grade is Grade II, and most patients have left-sided involvement. Most patients did not receive chemotherapy or hormone therapy, but a majority received radiotherapy. The data show that breast cancer in this population is more common in middle-aged women, with an equal proportion having no family history of the disease.

Table (2): Estimate Restricted Mean Survival Time (RMST), probability and months

Timepoint (τ)	RMST (Months)	RMST (Probability)	Standard Error	Lower C. I	Upper C. I
10	8.4	0.2486	0.07	8.2	8.65
20	19.3	0.1375	0.085	19	19.6
30	28.8	0.2526	0.148	28.3	29.3

45	42.9	0.1981	0.246	42.1	43.8
60	55.7	0.031	0.279	55.1	58.3
90	82.8	0.0147	0.658	80.7	85
125	85.5	0.0482	1.284	102.3	110.7

The table above shows the distribution of time was Weibull, and the estimated value of the parameters is:

$\alpha = 0.90583$ (Shape parameter), $\beta = 31.9119$ (Scale parameter),

$\tau = 10\,, 20\,, 30\,, 45\,, 60\,, 90\,, 125$

The Restricted Mean Survival Time (RMST) provides estimates of average survival times up to a predefined time point (denoted as τ), along with associated probabilities, standard errors, and confidence intervals.

Time point (τ): The specific time points (in months) at which the RMST is calculated. The time points are 10, 20, 30, 45, 60, 90, and 125 months. These time points represent the intervals at which survival is being evaluated.

RMST (Months): This provides the estimated average survival time (in months) up to the specified time point (τ). At 10 months, the RMST is 8.4 months, which means that, on average, patients survived 8.4 months within the first 10 months. At 125 months, the RMST is 85.5 months, indicating that, on average, patients survived 85.5 months within the first 125 months.

RMST (Probability): This column represents the probability associated with the RMST at each time point. It indicates the likelihood of survival up to that time point. For example, at 10 months, the probability is 0.2486 (24.86%), meaning there is a 24.86% chance of survival up to 10 months.

At 125 months, the probability drops to 0.0482 (4.82%), indicating a much lower likelihood of survival up to 125 months. The standard error of the RMST estimate measures the variability or uncertainty in the RMST estimate. A smaller standard error indicates more precision in the forecast. For example, at 10 months, the standard error is 0.070, which is relatively small, indicating a precise estimate. At 125 months, the standard error is 1.284, which is much larger, reflecting greater uncertainty in the forecast.

The lower and upper bounds of the 99% confidence interval for the RMST. The confidence interval gives a range within which the true RMST is expected to lie with 99% confidence.

When timepoint(τ) equals 10 months, the RMST is estimated to be between 8.2 and 8.65 months. When timepoint(τ) is 125 months, the RMST is estimated to be between 102.3 and 110.7 months.

Table (3): Comparison of Survival Curves

Log-rank (Mantel-Cox) test (conservative)	
Chi-square	2877
Df	5
P value	<0.0001
Are the survival curves sig different?	Yes
Log-rank test for trend (recommended)	
Chi-square	316.9
Df	1
P value	<0.0001
Sig. trend?	Yes
Gehan-Breslow-Wilcoxon test	
Chi-square	2255
Df	5
P value	<0.0001
Are the survival curves sig different?	Yes

The table (3) shows that survival analysis is often used in medical research to compare the time-to-event (e.g., death, recurrence) among different groups, such as treatment groups or patient subpopulations. To compare survival curves across treatment groups or subcategories (e.g., surgery, chemotherapy, radiation, hormone, Isotope, Target)

1. **Log-rank (Mantel-Cox) Test**

Chi-square statistic: 2877 (a very high value indicating significant between-group differences). Degrees of freedom (df): 5 (six groups compared). P-value: <0.0001 (indicating the results are highly

statistically significant).

The test compares the observed vs. expected number of events (e.g., deaths) at each time point across groups. Assumes proportional hazards (i.e., the ratio of hazard rates is constant over time).

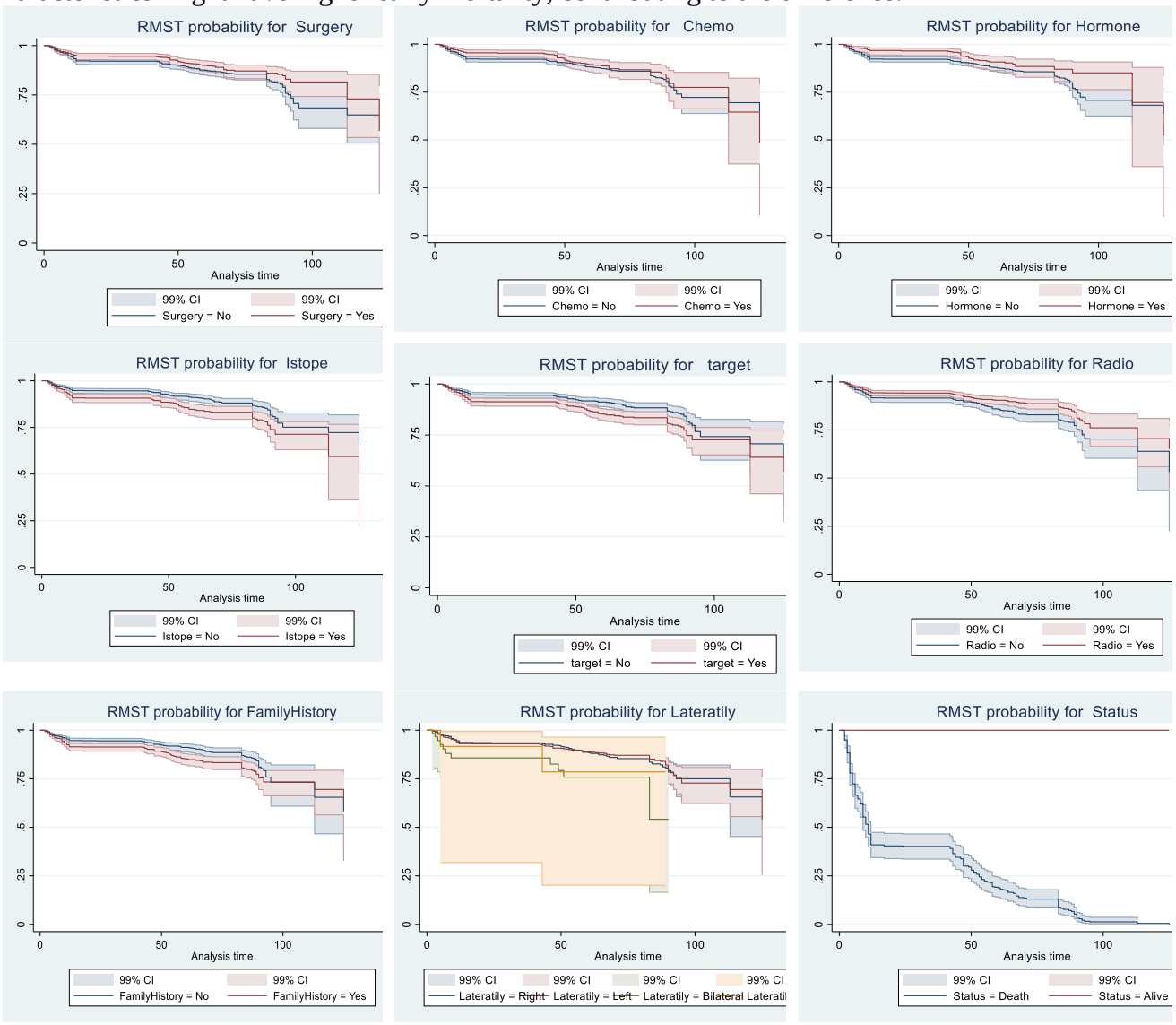
2. **Log-rank Test for Trend (Recommended)**

This test is a modified version of the log-rank test, focusing on detecting trends across ordered groups: Chi-square statistic: 316.9, degrees of freedom (df): 1, P-value: <0.0001.

Highly sensitive to ordered differences among groups, making it more precise when a trend is expected. This test confirms a significant trend in survival outcomes, suggesting a systematic improvement or worsening of survival across groups. For example, a linear relationship between treatment intensity and survival.

3. **Gehan-Breslow-Wilcoxon Test**

This test is similar to the log-rank test but places greater emphasis on early survival differences: Chi-square statistic: 2255, Degrees of freedom (df): 5, P-value: <0.0001. The significant result indicates that group survival differences are pronounced early in the follow-up period. For example, a treatment may have rapid early benefits but no long-term survival advantage. Groups with severe baseline characteristics might have higher early mortality, contributing to the difference.



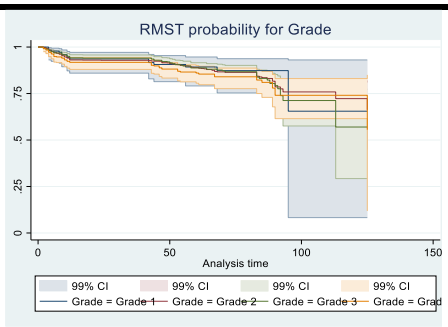


Figure 1: RMST probability to death in each treatment

In Figure (1) RMST probability of survival time to death within each treatment group (Status, surgery, radio, chemo, hormone, isotope, target, family history, laterality, and grade) contains a series of fixed images that visually represent the Restricted Mean Survival Time (RMST) probability of survival time to death across different treatment groups. It is divided into multiple panels, each likely corresponding to a specific treatment or variable group. While the actual images are not provided in the text, the description and context allow us to infer the content and purpose of these figures.

The figures are designed to visually compare the RMST probability of survival across various treatment groups and other explanatory variables. The statistics likely show how different treatments or factors (e.g., surgery, chemotherapy, radiotherapy, etc.) influence survival outcomes.

The figure is divided into multiple panels, each containing a graph or plot. Each panel likely shows RMST probability curves for different subgroups within a specific treatment or variable. The x-axis of each plot likely represents time (e.g., in months), while the y-axis represents the probability of survival. The curves in each panel likely compare the survival probabilities for different categories within a treatment group (e.g., patients who received surgery vs. those who did not).

If the survival curve for patients who received hormone therapy is higher than for those who did not, it suggests that hormone therapy may improve survival.

The figures also examine the impact of non-treatment variables, such as family history, laterality, and grade. For example, patients with a family history of the condition may have different survival probabilities than those without. The grade of the condition (e.g., Grade I vs. Grade IV) may significantly influence survival outcomes.

Table (4): Cumulative Incidence Probability for each Time.

Time	Number at Risk	Cumulative Incidence	Standard Error	95% Confidence Interval	
				Lower	Upper
2	13971	0.038	0.002	0.036	0.039
3	13091	0.132	0.001	0.13	0.134
4	11494	0.195	0.002	0.192	0.198
.	.	.	.	.	.
.	.	.	.	.	.
.	.	.	.	.	.
59	3418	0.677	0.001	0.673	0.68
60	3954	0.681	0.011	0.678	0.685
61	3241	0.691	0.007	0.687	0.694
.	.	.	.	.	.
.	.	.	.	.	.
.	.	.	.	.	.
95	195	0.901	0.041	0.898	0.903
113	149	0.904	0.001	0.902	0.906
125	57	0.908	0.012	0.906	0.91

The table (4) presents the Cumulative Incidence Function (CIF) for an event of interest over time. The CIF estimates the probability of the event occurring by a specific time point,

1. **Time**: The cumulative probability of the event of interest occurring by that time point. Represents the time points (e.g., days, months, or years) at which the cumulative incidence is calculated. At time 2, the cumulative incidence is 0.038 (or 3.8%). This calculates the CIF by ascending; at time

125, the cumulative incidence is 0.908, which increases over time as more events occur.

2. **Number at Risk**: The number of individuals who are still at risk of experiencing the event of interest at the beginning of the time interval. at time **2**, there are **13,971** individuals at risk.
3. **Standard Error**: Measures the uncertainty or variability in the estimate of the cumulative incidence. A more minor standard error indicates greater precision in the forecast.

Table (5): Cumulative Incidence Model

Factors	Estimator	SE	Z	P-Value	Confidence Interval	
					lower	upper
Intercept	3.036318	0.004205	-1.73	0.756	2.587456	3.587462
Gender	0.01204	0.471086	-0.06	0.301	0.004404	1.000969
Grade	-0.01908	0.07109	1.02	0.047	-0.3746	2.51289
Laterality	-0.03771	0.119876	1.5	0.024	-1.939159	1.218616
Surgery	-0.04643	0.124814	-0.79	0.001	-0.9542254	1.427292
Chemo	-0.01559	0.185761	0.73	0.004	-0.992642	1.052471
Radio	0.038368	0.095599	-1.72	0.003	0.0083061	1.088726
Hormone	0.046475	0.12601	-1.91	0.238	0.0079179	1.036307
Isotope	0.002354	0.211401	0.2	0.776	0.0006288	1.081302
Target	0.001311	0.272497	0.97	0.065	0.0002496	1.083464
Family History	0.00848	0.190305	0.72	0	0.000794	1.021735
Occupation	0.006001	0.00754	0.76	0.914	0.000442	1.052471

Most factors in the model, including **gender**, **Laterality**, **hormone therapy, isotope therapy, targeted therapy, and occupation,** do not show statistically significant results when comparing p-value with alpha ($\alpha = 0.05$) associations with the outcome. However**, Surgery, Chemo, Radio**, **grade and family history** are identified as significant predictors, suggesting their potential importance in the clinical context. From the Cumulative Incidence Model Equation (8)

$$\begin{aligned}\phi\big(F_1(t|z)\big) = 3.036 * \exp\,(&0.0.012gender - 0.019Grade - 0.037\text{Lateratily}\\ &- 0.046\text{Surgery} - 0.015\text{Chemo} + 0.038\text{Radio} + 0.046\text{Hormone}\\ &+ 0.002\text{Istope} + 0.001\text{target} + 0.008\text{FamilyHistory}\\ &+ 0.006\ \text{Occupation}\end{aligned} \quad (14)$$

Table (6): test compares K groups while accounting for Cumulative Incidence Function

Gray's Test				
	Statistic value	DF	P-Value	Reject H0
Chi-Square	329.754	2	0	
Pepe and Mori's Test	338.369	2	0	at α = 0.01and 0.05
Log-rank Test	278.136	2	0	

Gray's test is designed explicitly for comparing cumulative incidence functions (CIF) in the presence of competing risks.

Null Hypothesis (H0): The CIFs for the event of interest are equal across all K groups. Alternative Hypothesis (H1): The CIFs for the event of interest are not equal across all K groups. The null hypothesis is rejected at α = 0.05, meaning there are significant differences in the CIFs between the groups. This suggests that there are substantial differences in the CIFs between the groups, even after accounting for competing risks.

12. Conclusion

The following conclusions have been reached after analysing the data on breast cancer, and based on the results of the practical part:

1. The study recognised significant differences in survival based on factors such as age, gender, tumour grade, and family history. These findings underscore the heterogeneity of breast cancer and the importance of personalised treatment strategies.
2. The use of RMST, along with survival tests, the Log-rank and Gehan-Wilcoxon tests, provided a comprehensive view of survival differences across treatment groups. The results confirmed

statistically significant differences in survival curves, particularly in early-stage survival, emphasising the importance of timely interventions.

3. RMST offers a clinically meaningful summary of survival outcomes, allowing healthcare providers to communicate expected survival durations more effectively to patients. It can reduce patient anxiety and improve decision-making regarding treatment options.
4. The study found that patients who underwent surgery or received radiotherapy had significantly higher survival probabilities compared to those who did not. Additionally, factors such as family history, tumour grade, and age significantly influenced survival outcomes, underscoring the heterogeneity of breast cancer and the need for personalised treatment strategies.
5. The study's Cumulative Incidence Function results show that, at a level of a = 0.05, the most common factors affecting breast cancer are surgery, chemotherapy, radiotherapy, grade, and family history.

13. Supplementary material

(None).

14. Author's Contributions

Sami Ali Obed: Designed the research. Kurdistan Ibrahim Mawlood.: Writing and editing.

15. Funding

(None).

16. Data availability statement

I collected my dataset from the Ministry of Health at Rizgary Hospital centre for Cancer Diseases, based on daily records of cancer patients.

17. Acknowledgements

The authors would like to thank the Ministry of Health and the General Directorate of Health–Erbil for providing medical data. Special appreciation is also extended to Rizgary Hospital Centre for Cancer Diseases for supplying the breast cancer dataset. The data included 4,420 cases of breast cancer patients, collected over five years and eight months, from January 1, 2019, to August 31, 2024.

18. Conflict of interest

The authors declare no conflict of interest.

References

[1] Abulizi, Xianmixikemaier, Heather J. Ribaudo, and Philippe Flandre. 2019. The Use of the Restricted Mean Survival Time as a Treatment Measure in HIV/AIDS Clinical Trial: Reanalysis of the ACTG A5257 Trial.

[2] A'Hern, Roger P. 2016. 'Restricted Mean Survival Time: An Obligatory End Point for Time-to-Event Analysis in Cancer Trials?' Journal of Clinical Oncology 34(28):3474–76. https://doi: 10.1200/JCO.2016.67.8045.

[3] Calkins, Keri L., Chelsea E. Canan, Richard D. Moore, Catherine R. Lesko, and Bryan Lau. 2018. 'An Application of Restricted Mean Survival Time in a Competing Risks Setting: Comparing Time to ART Initiation by Injection Drug Use'. BMC Medical Research Methodology 18(1). https://doi: 10.1186/s12874-018-0484-z.

[4] Chiumente, Marco, Daniele Mengato, and Andrea Messori. 2020. 'Tisagenlecleucel in Non-Hodgkin Lymphoma: The Restricted Mean Survival Time as a Tool for Estimating Progression-Free Life Expectancy Better than the Median'. Acta Haematologica 143(6):598–99.

[5] He, Yifan, and Yong Zhou. 2020. 'Nonparametric and Semiparametric Estimators of Restricted Mean Survival Time under Length-Biased Sampling'. Lifetime Data Analysis 26(4):761–88. https://doi: 10.1007/s10985-020-09498-x.

[6] Hajer, F. T., & Ibrahim, W. S. (2029). Use the logistic regression model to find the most critical factors affecting lung cancer in Iraq in 2017, Journal of Administration and Economics, 44(121).

[7] Ibrahim, W. S., & Jaber. A. K. A. (2023). A study of patient with bladder cancer using the survival function of the new expanded adapter Whipple distribution. Journal of Administration and Economics, 48(139). https://doi.org/10.31272/jae.i139.1090

[8] Inam, D. R., Noaman, A., Huda, ∗ A L, Abd, A., & ∗∗ A.-A. (2019). Comparison of Classical and Bayesian methods to Estimate the shape parameter and Reliability function in Burr type X or two param4eter of exponential Rayleigh distribution under different Loss function. https://doi.org/10.31272/JAE.42.2019.11

[9] Jackson, Richard J. 2015. Some Statistical Methods for the Analysis of Survival Data in Cancer Clinical Trials.

[10] Mawlood, Kurdistan I., & Obed, Sami A. (2019). QALAAI ZANIST SCIENTIFIC JOURNAL Study and Analysis of the Chest Cancer Data Using Survival Models ARTICLE INFO ABSTRACT. A Scientific Quarterly Refereed Journal Issued by Lebanese French University-Erbil, 4. https://doi.org/10.25212/lfu.qzj.4.2.24

[11] Lyu, Jingjing, Yawen Hou, and Zheng Chen. 2020a. 'The Use of Restricted Mean Time Lost under Competing Risks Data'. BMC Medical Research Methodology 20(1). https://doi: 10.1186/s12874-020-01040-9.

[12] Messori, Andrea, Vera Damuzzo, Laura Agnoletto, Luca Leonardi, Marco Chiumente, and Daniele Mengato. 2020. 'A Model-Independent Method to Determine Restricted Mean Survival Time in the Analysis of Survival Curves'. SN Comprehensive Clinical Medicine 2(1):66–68. https://doi: 10.1007/s42399-019-00199-7.

[13] Trinquart, Ludovic, Justine Jacot, Sarah C. Conner, and Raphaël Porcher. 2016. 'Comparison of Treatment Effects Measured by the Hazard Ratio and by the Ratio of Restricted Mean Survival Times in Oncology Randomized Controlled Trials'. Journal of Clinical Oncology 34(15):1813–19. https://doi: 10.1200/JCO.2015.64.2488.

[14] Zucker, David M. 1998. 'Restricted Mean Life with Covariates: Modification and Extension of a Useful Survival Analysis Method'. Journal of the American Statistical Association 93(442):702–9. https://doi: 10.1080/01621459.1998.10473722.

https://doi.org/10.31272/jae.i149.1395
https://admics.uomustansiriyah.edu.iq/index.php/admeco
P-ISSN: 1813–6729 E-ISSN: 2707-1359

متوسط وقت البقاء المقيد يقترب من الآثار المترتبة على بيانات سرطان الثدي

سامي على عبيد
جامعة صلاح الدين / كلية الادارة و الاقتصاد/ قسم الاحصاء و المعلوماتية/ اربيل
Email: sami.obed@su.edu.krd, ORCID ID:/ https://orcid.org/0000-0002-2866-5886
كوردستان ابراهيم مولود
جامعة صلاح الدين / كلية الادارة و الاقتصاد/ قسم الاحصاء و المعلوماتية/ أربيل
Email: kurdistan.mawlood@su.edu.krd , ORCID ID :/ https://orcid.org/0000-0002-1612-1996

المستخلص

يقيم هذا البحث فائدة متوسط وقت البقاء المقيد (RMST)كبديل قابل للتطبيق لتقنيات تحليل البقاء التقليدية، بما في ذلك اختبار كابلان ماير ونموذج المخاطر النسبية كوكس، خاصة عندما لا يتم استيفاء افتراض المخاطر النسبية . تقدم RMSTنهجا بديلا لاستخدام نسب الخطر في مثل هذه المواقف .هدفت هذه الدراسة إلى تقييم تأثير العوامل التنبؤية على البقاء على قيد الحياة من سرطان الثدي يقدم متوسط وقت البقاء المقيد (RMST)مقياسا قابلا للتفسير سريريا لمتوسط وقت البقاء على قيد الحياة خلال فترة محددة مما يوفر فهما أكثر تفصيلا لتأثيرات العلاج ونتائج المرضى .بالإضافة إلى ذلك، استخدمت الدراسة نموذج دالة الوقوع التراكمي (CIF)وتقديرات الاحتمالات لكل نقطة زمنية لتحليل بيانات البقاء على قيد الحياة بشكل أكبر.

باستخدام بيانات من 4,420مريضا بسرطان الثدي في مستشفى رزكاري، يحدد هذا البحث تأثير العوامل النذيرية المختلفة على نتائج البقاء على قيد الحياة .توضح الدراسة أن RMSTيلتقط بشكل فعال الاختلافات في احتمالات البقاء على قيد الحياة عبر مجموعات العلاج، حيث يظهر المرضى الذين يخضعون لعملية جراحية أو يتلقون العلاج الإشعاعي معدلات بقاء أعلى بكثير من أولئك الذين لم يفعلوا ذلك .بالإضافة إلى ذلك، يكشف التحليل أن عوامل مثل التاريخ العائلي ودرجة الورم تؤثر بشكل كبير على نتائج البقاء على قيد الحياة، مما يسلط الضوء على عدم تجانس سرطان الثدي .تؤكد المقارنات الإحصائية باستخدام RMST، جنبا إلى جنب مع الاختبارات التقليدية مثل اختبارات Log-rank وGehan-Wilcoxon، اختلافات كبيرة في منحنيات البقاء على قيد الحياة ، لا سيما في البقاء على قيد الحياة في المراحل المبكرة ، مما يؤكد أهمية التدخلات في الوقت المناسب.

معلومات البحث

تواريخ البحث:

تاريخ تقديم البحث: 16 / 03 / 2025
تاريخ قبول البحث: 24 / 07 / 2025
تاريخ نشر الكتروني: 01 / 09 / 2025
عدد صفحات البحث 30 - 41

الكلمات المفتاحية:

RMST، سرطان الثدي ، تحليل البقاء على قيد الحياة ، منحنى البقاء على قيد الحياة ، CIF:

المراسلة:

أسم الباحث: كوردستان ابراهيم مولود

Email: kurdistan.mawlood@su.edu.krd