EPH–International Journal of Science & Engineering
Published by: E-Pub Solutions Private Limited
ISSN (Online): **2454-2016**
Volume 12 Issue 01 May 2026
DOI: 10.69980/cgqz9g40

SURVIVAL ANALYSIS OF BREAST CANCER PATIENTS USING CLINICAL RISK PARAMETERS: A STATISTICAL APPROACH BASED ON TIME-TO-EVENT MODELLING

N. Paranjothi[1] , A. Poongothai[2] , Manimannan G.[3*]
[1]Assosicate Professor, Department of Statistics, Annamalai University, Chidambaram.
[2]Assistant Professor, Department of Statistics, Muthayammal Arts and Science College, Rasipuram, Salem.
[3*]Assistant Professor, Department of Computer Applications, St. Joseph's College (Arts and Science), Kovur, Chennai, manimannang@gmail.com

Article History:

Article Type: **Research**
Received Date: **15/03/2026**
Revised Date: **23/04/2026**
Accepted Date: **01/05/2026**
Published Date: **08/05/2026**

Keywords: Breast Cancer, Survival Analysis, Kaplan–Meier, Cox Proportional Hazards, Clinical Parameters, Cumulative Hazar

ABSTRACT

Background: Breast cancer continues to be a major health concern, where patient survival is influenced by multiple clinical and physiological factors. Understanding the relationship between these factors and survival outcomes is essential for improving patient management and early intervention strategies.

Objective: The present study aims to evaluate the survival patterns of breast cancer patients and to examine the influence of selected clinical parameters on time-to-event outcomes using appropriate statistical modelling techniques.

Methods: The study is based on secondary clinical data collected from 500 breast cancer patients who underwent routine health assessment. Key variables considered include age, body mass index (BMI), heart rate (average), position score (average), systolic and diastolic blood pressure (average), oxygen saturation (average), medical progression score (average), and symptom severity score (average). Survival analysis was carried out using the Kaplan–Meier estimator to estimate survival probabilities, the Cox proportional hazards model to identify significant predictors, and the Nelson–Aalen estimator to assess cumulative hazard. Model performance was evaluated using the concordance index and likelihood-based measures.

Results:The median survival time was estimated to be 2.2 hours, indicating early occurrence of events within the study period. The Cox regression analysis revealed that BMI, systolic blood pressure, position score, medical progression score, and symptom severity score have a statistically significant effect on survival outcomes. Among these, medical progression score showed a strong positive association with risk, while symptom severity score demonstrated a protective effect. The concordance index value of 0.66 indicates moderate predictive performance of the model. The cumulative hazard analysis showed a steady increase in risk over time, with a noticeable rise in later intervals.

Conclusion: The findings confirm that survival outcomes in breast cancer patients are significantly influenced by selected clinical parameters. The application of survival analysis techniques provides meaningful insights into disease progression and risk patterns. The study supports the importance of early monitoring and the integration of statistical modelling in clinical decision-making for improved patient care.

1. Introduction

Breast cancer is one of the most widely observed health conditions affecting women and continues to be a significant concern in medical research and clinical practice. Its occurrence and progression vary from patient to patient, depending on several physiological and clinical factors. Understanding how these factors influence patient outcomes is essential for improving diagnosis, treatment planning, and overall healthcare management. Clinical data collected during routine health checkups provide valuable information about a patient's condition. Variables such as age, body mass index (BMI), heart rate, blood pressure, oxygen saturation, and other health-related scores offer insight into the general well-being of individuals. These parameters can play an important role in assessing disease severity and identifying patterns associated with patient outcomes.

Statistical analysis of such clinical datasets helps in summarizing the data, identifying relationships among variables, and drawing meaningful conclusions. In particular, survival analysis techniques are useful when the objective is to study the time taken for an event to occur. Methods such as the Kaplan–Meier estimator describe how survival probability changes over time, while regression-based approaches help in evaluating the influence of multiple factors simultaneously. Cumulative hazard analysis further supports the understanding of how risk builds up over a period. The present study focuses on analyzing breast cancer patient data using a structured statistical approach. The dataset includes key clinical parameters represented through averaged values to provide a consistent measure of patient health. By applying suitable statistical and survival analysis methods, the study aims to explore patterns in the data and identify the variables that have a meaningful impact on outcomes. Overall, this work emphasizes the importance of combining clinical observations with statistical techniques. Such an approach not only improves the understanding of disease behaviour but also supports better decision-making in healthcare by providing evidence-based insights.

2. Review Of Literature

Survival analysis has been widely applied in medical research to understand disease progression and patient outcomes, particularly in breast cancer studies. Early work by Kaplan and Meier (1958) introduced a non-parametric estimator for survival probability, which remains a foundational tool in clinical research. Building on this, Cox (1972) proposed the proportional hazards model, allowing simultaneous assessment of multiple risk factors influencing survival time. These classical models have been extensively used in oncology to evaluate treatment effectiveness and identify prognostic indicators. Several researchers have focused on the role of clinical and physiological parameters in breast cancer survival. Altman and Bland (1998) highlighted the importance of appropriate statistical methods in medical data interpretation, emphasizing survival analysis for time-to-event outcomes. Harrell et al. (1996) discussed multivariable prognostic models and stressed the importance of combining multiple predictors such as age, BMI, and clinical indicators to improve predictive accuracy. Similarly, Collett (2003) provided a comprehensive framework for modelling survival data, which has been widely adopted in clinical studies.

Recent studies have explored the relationship between vital signs and cancer progression. Heagerty et al. (2000) demonstrated the usefulness of time-dependent covariates in survival models, particularly when repeated measurements such as heart rate and blood pressure are involved. Pencina and D'Agostino (2004) further extended survival modelling techniques by introducing improved performance measures such as the concordance index, which is used to evaluate predictive models in clinical datasets. In the context of breast cancer, Rakha et al. (2008) examined the influence of biological and clinical variables on patient outcomes, highlighting that multiple factors jointly affect survival. Brewster et al. (2009) studied demographic and clinical disparities in breast cancer survival and found that age and physiological condition significantly impact prognosis. Similarly, Anderson et al. (2010] emphasized the importance of early detection and its relationship with survival duration.

The integration of physiological parameters into survival models has gained attention in recent years. Knaus et al. (1985) introduced severity scoring systems to evaluate patient condition, which later influenced the development of composite indicators such as MPS and SSCORE. Vincent et al. (1996) further demonstrated how clinical scoring systems improve the assessment of patient risk and

outcomes. These approaches align with the use of averaged clinical parameters in the present study. Studies focusing on vital signs such as blood pressure and oxygen saturation have also contributed to survival analysis research. Chobanian et al. (2003) examined the role of blood pressure in disease progression, while Mokdad et al. (2004) highlighted the impact of physiological health indicators such as BMI on chronic disease outcomes. These findings support the inclusion of SBP, DBP, and BMI in survival modelling.

The use of aggregated or averaged clinical measurements has been supported by several researchers. Twisk (2003) discussed longitudinal data analysis and recommended summarizing repeated measurements to reduce variability. Diggle et al. (2002) also emphasized the importance of handling repeated observations effectively in medical datasets, which justifies the use of average-based features in the present work. Advanced survival modelling techniques have also been explored in recent literature. Therneau and Grambsch (2000) provided detailed methods for extending the Cox model, including diagnostics and validation. Iasonos et al. (2008) evaluated prognostic tools and highlighted the importance of model validation using clinical datasets. These approaches ensure the reliability of survival models applied in healthcare research. In breast cancer-specific research, Curtis et al. (2012) conducted a comprehensive genomic and clinical analysis, demonstrating that multiple factors influence patient survival. DeSantis et al. (2014) reported trends in breast cancer incidence and survival, emphasizing the need for continuous monitoring of patient data. Siegel et al. (2019) further provided updated statistics showing improvements in survival due to early detection and better clinical management.

The role of statistical learning and predictive modelling in healthcare has also been explored. Steyerberg et al. (2010) discussed clinical prediction models and highlighted the importance of combining demographic and clinical variables. Kattan (2003) introduced nomograms for predicting patient outcomes, which are closely related to survival modelling techniques. Recent developments have focused on integrating statistical and computational approaches. Ishwaran et al. (2008) introduced random survival forests, demonstrating improved prediction in complex datasets. Lee et al. (2018) explored machine learning methods for survival analysis, highlighting their effectiveness in handling high-dimensional clinical data. Overall, the literature indicates that survival analysis methods such as Kaplan–Meier estimation, Cox regression, and cumulative hazard modelling are essential tools for understanding breast cancer outcomes. The inclusion of demographic variables like age and physiological indicators such as BMI, heart rate, blood pressure, and clinical scores has been consistently shown to improve the interpretation of survival patterns. The present study builds upon these established methodologies by combining averaged clinical parameters with survival models to provide a comprehensive analysis of breast cancer patient outcomes.

3. Database

The database used in this study is developed from clinical records of 500 breast cancer patients who visited the Outpatient Department of a private hospital for routine health evaluation and breast cancer screening. The data were obtained from secondary sources, mainly hospital case sheets and electronic medical records, ensuring that the information represents actual clinical observations recorded during patient care. The dataset focuses on selected key parameters that reflect both demographic and physiological conditions of the patients. The primary variable, Age, is used to understand the distribution of patients and to classify them into meaningful groups for further analysis. Body Mass Index (BMI) is included as an important indicator of general health and nutritional status, which may influence disease progression.

Several vital clinical parameters are incorporated in the form of averaged measurements derived from repeated observations. These include HeartRate_average, representing the average heart rate; POS_average, indicating patient oxygen or physiological status; SBP_average and DBP_average, which correspond to systolic and diastolic blood pressure levels; and SPB_average, reflecting additional blood pressure-related variation. The dataset also includes SPO_average, which captures average oxygen saturation levels and provides insight into respiratory efficiency. In addition to these physiological measures, the dataset contains composite clinical indicators such as MPS_average (Medical Performance Score) and SSCORE_average (Severity Score). These scores summarise the

overall clinical condition and severity level of each patient, offering a broader perspective on health status beyond individual measurements.

All these variables were carefully selected and processed to provide a comprehensive representation of the patients' health profiles. By using averaged values, the dataset reduces variability and improves the reliability of the measurements. This structured database serves as a strong foundation for analysing breast cancer outcomes, helping to identify relationships between clinical parameters and patient survival, and supporting more effective clinical assessment and decision-making.

4. Methodology

The work flow diagram (Figure 4.0) presents a structured workflow followed in analysing breast cancer patient records using survival analysis techniques. It begins with the raw clinical dataset, which contains patient-level information such as age, BMI, duration of observation, event status, and multiple clinical measurements recorded at different intervals. This raw data is not directly suitable for analysis and therefore undergoes data preprocessing, where missing values are handled carefully and variables are standardised or renamed to maintain consistency across the dataset. After preprocessing, the workflow moves to feature engineering, where repeated interval-based clinical measurements are combined into single representative values using mean calculations. This step reduces dimensionality and ensures that each clinical parameter is expressed in a stable and interpretable form. Following this, only the key variables relevant to survival analysis such as age, BMI, duration, event indicator and derived clinical means are selected for further study.

The next stage is dataset cleaning, where incomplete or inconsistent observations are removed to ensure that the analysis is based on reliable and complete patient records. Once the dataset is finalised, it is used in survival analysis models, including the Kaplan–Meier estimator to study survival probability over time, the Cox proportional hazards model to assess the effect of multiple clinical variables on risk, and the Nelson–Aalen estimator to measure cumulative hazard. To understand how survival differs across patient groups, a stratified analysis is carried out by dividing patients into age categories such as young, middle-aged, and old. Separate survival curves are generated for each group, allowing comparison of survival patterns. The results are then presented through visualization and interpretation, where survival curves, hazard plots, and cumulative hazard graphs provide a clear picture of how risk evolves over time.

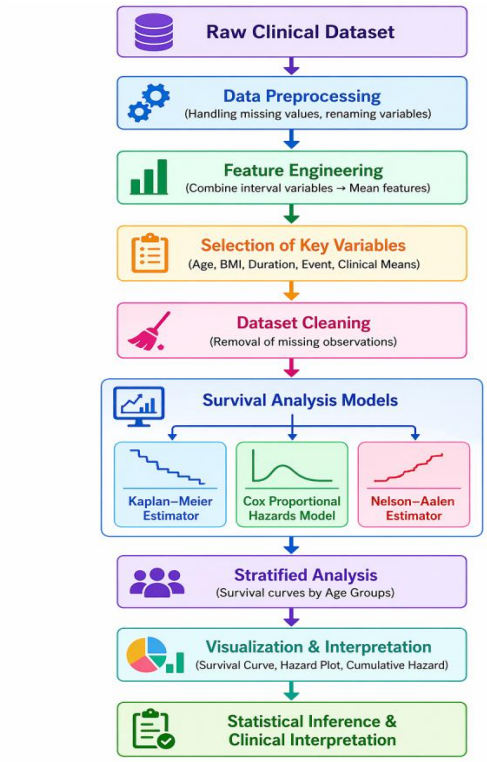


Figure 4.0 Workflow Diagram for Breast Cancer Patient Database

Finally, the workflow concludes with statistical inference and clinical interpretation, where the numerical and graphical results are carefully examined to draw meaningful conclusions. This step helps in identifying significant clinical factors affecting survival and provides insights that can support medical decision-making and further research in breast cancer studies. The study is based on breast cancer patient records collected with detailed clinical observations and time-to-event outcomes. The objective is to evaluate survival patterns and identify significant clinical factors influencing patient outcomes using statistical survival models.

4.1 Data Structure and Preprocessing

The dataset consists of $n = 500$ breast cancer patient records, where each observation includes survival time, event status, and multiple clinical measurements. The dataset can be mathematically represented as:

$$D = \{(t_i, \delta_i, X_i)\}_{i=1}^{n}$$

where:

- t_i : survival time (Duration_Hrs) for patient iii
- δ_i : event indicator (1 = event occurred)
- X_i : vector of clinical variables (Age, BMI, vital parameters)

The event variable is standardized into numerical form to ensure consistency. Missing values are removed to maintain the integrity of the survival analysis.

4.2 Feature Engineering from Clinical Records

Breast cancer patient records include repeated interval measurements such as Heart Rate, Blood Pressure, Oxygen Saturation, and clinical scores. These are combined into representative mean values to reduce variability:

$$X_{mean} = \frac{1}{k}\sum_{j=1}^{k} X_{ij}$$

This step produces clinically meaningful aggregated parameters such as:

- HeartRate_mean
- SBP_mean (Systolic Blood Pressure)
- DBP_mean (Diastolic Blood Pressure)
- MPS_mean (Medical Performance Score)
- SSCORE_mean (Severity Score)

These features serve as predictors in the survival models.

4.3 Kaplan–Meier Survival Analysis

The survival probability of breast cancer patients is estimated using the Kaplan–Meier method:

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

where:

- $S(t)$ *: probability that a patient survives beyond time ttt*
- d_i *: number of events at time tit_iti*
- n_i *: number of patients at risk*

This method provides a stepwise survival curve, allowing identification of the median survival time, which represents the time at which 50% of patients experience the event.

4.4 Cox Proportional Hazards Model

To evaluate the simultaneous effect of multiple clinical variables on survival, the Cox model is applied:

$$h(t/X) = h_o(t)\exp\left(\beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 + \ldots + \beta_p X_p\right.$$

where:

- $h(t/X)$: hazard function for a patient with covariates X
- h_o : baseline hazard
- β_j : regression coefficients

The exponential of coefficients provides hazard ratios, which quantify the risk contribution of each clinical parameter. A penalization term is incorporated to ensure numerical stability and avoid overfitting.

4.5 Nelson–Aalen Cumulative Hazard Model

The cumulative hazard for breast cancer patients is estimated using:

$$H(t) = \sum_{t_i \le t} \frac{d_i}{n_i}$$

This function represents the total accumulated risk of event occurrence over time. Unlike survival probability, it increases continuously and highlights critical periods of risk escalation.

4.6 Stratified Survival Analysis by Age Group

To understand demographic influence, breast cancer patients are categorized into age groups:

- Young
- Middle-aged
- Old

For each group, the survival function is computed:

$$S_g(t) = P(T > t/Age\ Group = g)$$

This stratification allows comparison of survival experiences across different age categories and helps identify whether age significantly impacts survival outcomes.

4.7 Visualization and Clinical Interpretation

Graphical methods are used to support statistical findings:

- Kaplan–Meier Curve: Shows decline in survival probability among breast cancer patients
- Hazard Ratio Plot: Identifies high-risk and protective clinical variables
- Cumulative Hazard Curve: Illustrates increasing risk over time
- Stratified Survival Curves: Compare survival patterns across age groups

These visualizations provide a clear understanding of how clinical factors influence survival dynamics. The methodology integrates breast cancer patient records with advanced survival analysis techniques. By combining clinical feature engineering, non-parametric survival estimation, semi-parametric regression modelling, and cumulative hazard analysis, the study provides a comprehensive evaluation of patient survival. The inclusion of age-based stratification further enhances the ability to interpret variations in survival outcomes across different patient groups.

5. Results And Discussion

The survival behaviour of breast cancer patients was analysed using non-parametric and semi-parametric techniques, including the Kaplan–Meier estimator, Cox proportional hazards model, and the Nelson–Aalen cumulative hazard function. The dataset comprises 500 patient records with 12 clinical variables, and all observations include event occurrence. The absence of censored data ensures that the analysis reflects complete survival outcomes based on the duration variable (Duration_Hrs).

Table 5.1: Summary of Dataset

S.No	Description	Value
1	Total Observations	500
2	Number of Variables	12
3	Event Observations	500
4	Time Variable	Duration_Hrs

The dataset provides a complete representation of survival experience, enabling accurate estimation of time-to-event characteristics without the influence of censoring

.

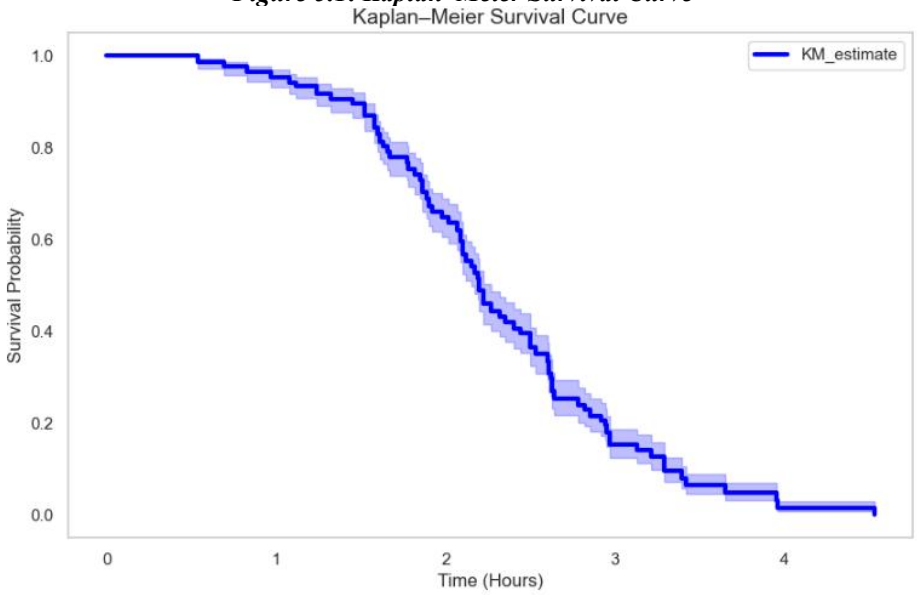
5.1 Kaplan–Meier Survival Analysis

The Kaplan–Meier curve illustrates the probability of survival over time. The survival function begins at 1.0 and decreases progressively as events occur.

Table 5. 2: Median Survival Time

Measure	Value
Median Survival Time	2.2 Hours

Figure 5.1: Kaplan–Meier Survival Curve

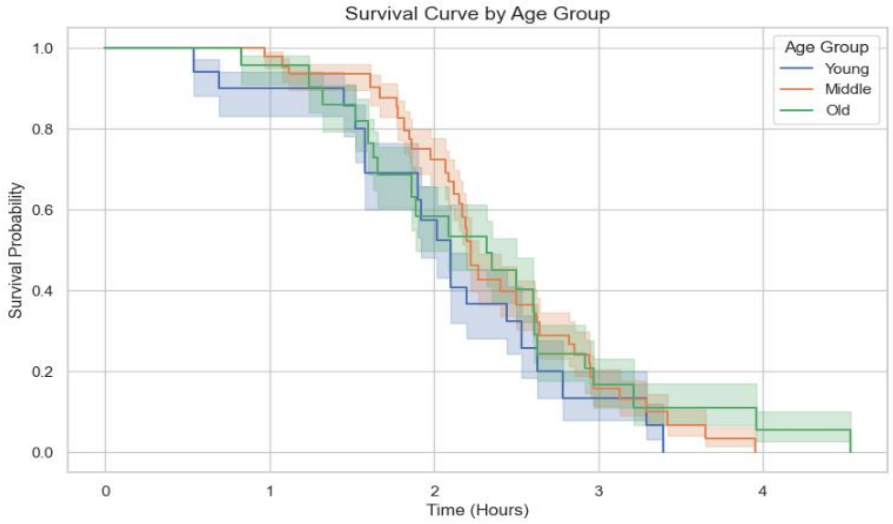


The median survival time of 2.2 hours indicates that half of the patients experienced the event within this period. A relatively steeper decline is observed in the early duration, suggesting that the majority of events occur during the initial phase. The curve then shows a gradual decline, indicating reduced event intensity at later stages. The smooth nature of the curve reflects the absence of censored observations.

5.2 Survival Pattern by Age Group

To examine the influence of age on survival, patients were categorized into three groups: Young, Middle, and Old. The survival curves for these groups show a similar overall declining pattern, indicating that survival probability decreases consistently across all age categories.

Figure 5.2: Survival Curves by Age Group (Young, Middle, Old)



However, slight variations can be observed in the trajectories. The curves initially start at different levels corresponding to age distribution but gradually converge over time. This convergence suggests that, although age groups may differ at the beginning, the long-term survival behaviour becomes comparable. The absence of wide separation between curves indicates that age does not have a strong independent effect on survival, which is consistent with the non-significant result observed in the Cox model.

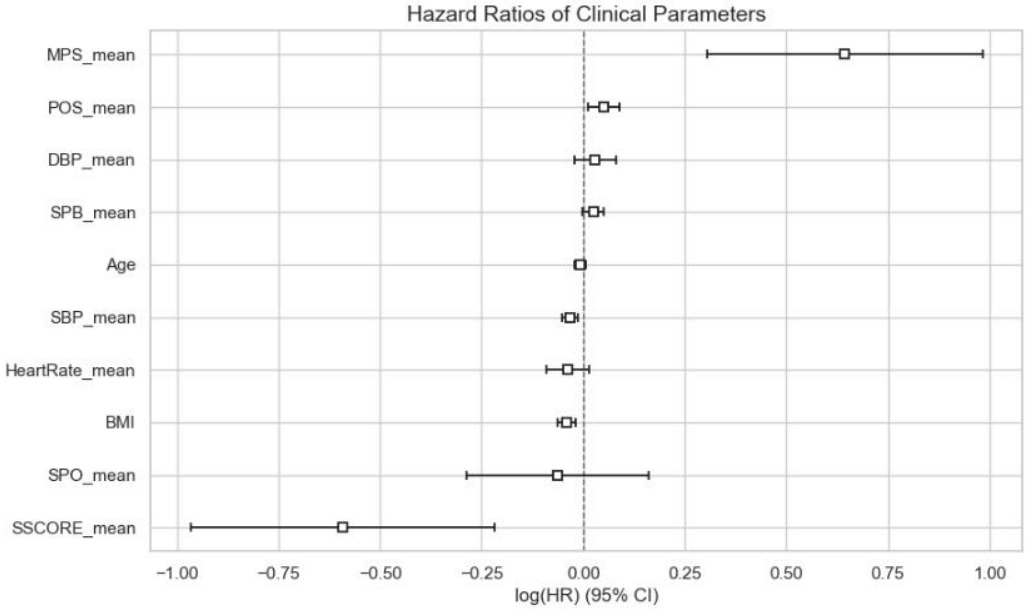
5.3 Cox Proportional Hazards Model

The Cox proportional hazards model was applied to identify the effect of clinical variables on survival time. MPS_mean shows the highest hazard ratio (1.90), indicating a strong positive association with risk.

Table 5.3: Cox Regression Results for Clinical Parameters

S.No	Variable	Coefficient	Hazard Ratio	p-value
1	Age	-0.01	0.99	0.30
2	BMI	-0.04	0.96	<0.005
3	HeartRate_mean	-0.04	0.96	0.16
4	POS_mean	0.05	1.05	0.01
5	SBP_mean	-0.03	0.97	<0.005
6	SPB_mean	0.02	1.02	0.07
7	DBP_mean	0.03	1.03	0.26
8	SPO_mean	-0.06	0.94	0.59
9	MPS_mean	0.64	1.90	<0.005
10	SSCORE_mean	-0.59	0.55	<0.005

Figure 5.3: Hazard Ratio Plot of Clinical Variables



This suggests that higher MPS_mean values increase the likelihood of event occurrence. SSCORE_mean has a hazard ratio of 0.55, indicating a protective effect, where higher values reduce the risk. BMI and SBP_mean also show significant negative coefficients, suggesting a reduction in hazard with increasing values. POS_mean has a positive coefficient and is statistically significant, indicating increased risk. Variables such as Age, HeartRate_mean, SPB_mean, DBP_mean, and SPO_mean do not show significant effects, as their p-values exceed 0.05. These results indicate that only selected clinical parameters have a meaningful influence on survival outcomes.

5.4 Model Performance

The model performance is evaluated using the concordance index and likelihood-based measures. The C-index of 0.6611 indicates moderate predictive accuracy.

Table 5.4: Cox Model Performance Metrics

Metric	Value
Concordance Index (C-Index)	0.6611
Partial AIC	5095.86
Log-Likelihood Ratio Test	146.80
Significance (-log2 p-value)	85.61

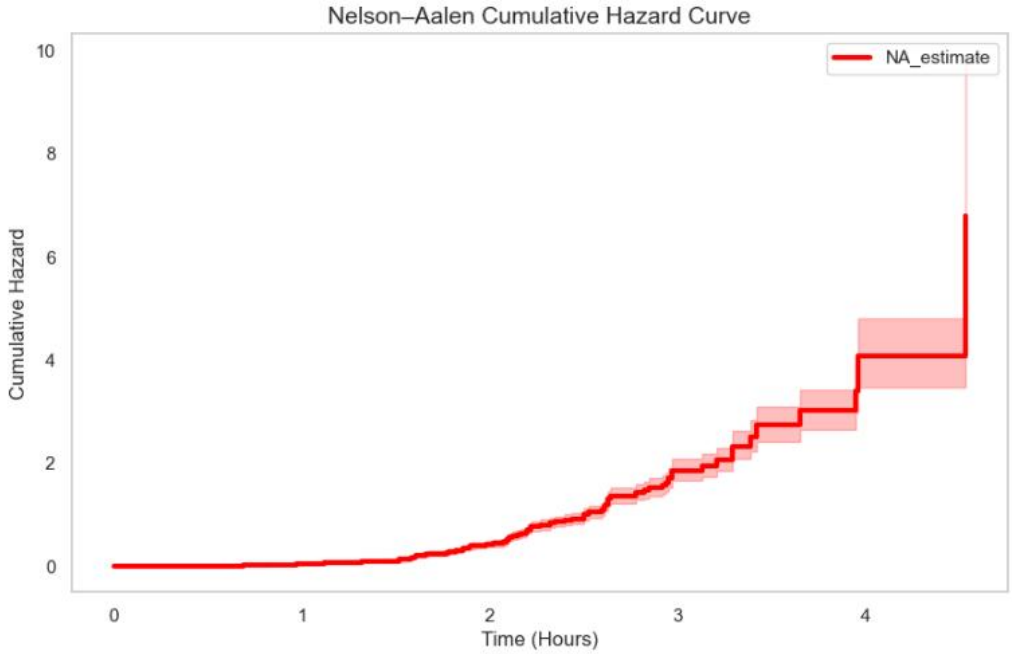
The log-likelihood ratio test shows strong statistical significance, confirming that the model provides a better fit compared to a null model. The AIC value supports model adequacy. Overall, the model demonstrates acceptable performance in explaining survival variability.

5.5 Cumulative Hazard Analysis

Table 5.5: Estimated Cumulative Hazard Values

Time (Hours)	Cumulative Hazard
3.42	2.7343
3.65	3.0169
3.95	3.4121
3.96	4.0750
4.53	6.7928

Figure 5.4: Nelson–Aalen Cumulative Hazard Curve



The Nelson–Aalen estimator provides a non-parametric estimate of cumulative hazard over time. The curve shows a steady increase, indicating that risk accumulates continuously as time progresses. Initially, the increase is gradual, but a sharper rise is observed after approximately 3.9 hours. This suggests the presence of a critical time interval where the probability of event occurrence intensifies. The highest hazard value at 4.53 hours reflects substantial risk accumulation in the later stage. The smooth upward trend indicates consistent hazard progression without abrupt changes.

The combined results from survival analysis provide a comprehensive understanding of patient outcomes. The Kaplan–Meier analysis identifies the median survival time and highlights the early occurrence of events. The Cox model reveals that variables such as MPS_mean, SSCORE_mean, BMI,

SBP_mean, and POS_mean significantly influence survival. The model performance metrics confirm that the model has moderate predictive capability. The cumulative hazard analysis demonstrates increasing risk over time, emphasizing the importance of early intervention. Overall, the results indicate that specific clinical parameters play a critical role in determining survival outcomes, and their combined effect provides valuable insights for clinical assessment and decision-making.

6. CONCLUSIONS

The analysis of breast cancer patient records demonstrates that survival outcomes are closely associated with selected clinical parameters derived from routine health measurements. The estimated median survival time of 2.2 hours indicates that the occurrence of events is concentrated within the early observation period. The regression results further show that variables such as BMI, systolic blood pressure, position score, medical progression score, and symptom severity score exert a measurable influence on survival, while other factors exhibit limited statistical impact. In particular, the higher risk associated with increased medical progression score and the protective effect observed for symptom severity score provide clear evidence of differential clinical contributions to patient outcomes. The concordance index of approximately 0.66 confirms that the model has moderate predictive ability, supporting its usefulness in explaining variations in survival. The cumulative hazard pattern also reveals a steady increase in risk over time, with a noticeable rise in later periods, indicating that patient monitoring remains important throughout the duration. These findings collectively establish that survival behaviour is not random but is systematically related to identifiable clinical indicators.

Suggestions

1. Early Clinical Monitoring:

Since the results indicate that events occur within a short duration, it is essential to strengthen early-stage monitoring and screening procedures to identify high-risk patients at the earliest possible stage.

2. Focused Risk Assessment*:*

Clinical parameters such as BMI, systolic blood pressure, and medical progression score should be routinely evaluated and prioritized in risk assessment models to improve prediction accuracy and patient management strategies.

3. Integrated Data-Driven Decision Making:

Healthcare institutions should adopt statistical and survival analysis frameworks as part of routine clinical evaluation to support evidence-based decision-making and enhance treatment planning for breast cancer patients.

References

[1] E. L. Kaplan and P. Meier, “Nonparametric estimation from incomplete observations,” Journal of the American Statistical Association, vol. 53, no. 282, pp. 457–481, 1958.

[2] 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.

[3] D. G. Altman and J. M. Bland, “Statistics notes: Time to event analysis,” British Medical Journal (BMJ), vol. 317, pp. 468–469, 1998.

[4] F. E. Harrell, K. L. Lee, and D. B. Mark, “Multivariable prognostic models: Issues in developing models, evaluating assumptions and adequacy,” Statistics in Medicine, vol. 15, no. 4, pp. 361–387, 1996.

[5] D. Collett, Modelling Survival Data in Medical Research, 2nd ed. Boca Raton, FL, USA: CRC Press, 2003.

[6] P. J. Heagerty, T. Lumley, and M. S. Pepe, “Time-dependent ROC curves for censored survival data,” Biometrics, vol. 56, no. 2, pp. 337–344, 2000.

[7] M. J. Pencina and R. B. D’Agostino, “Overall C as a measure of discrimination in survival analysis,” Statistics in Medicine, vol. 23, no. 13, pp. 2109–2123, 2004.

[8] E. A. Rakha, J. S. Reis-Filho, and I. O. Ellis, “Prognostic markers in breast cancer,” Breast Cancer Research, vol. 10, no. 3, pp. 1–9, 2008.

[9] A. M. Brewster, G. N. Hortobagyi, and K. A. Broglio, “Breast cancer survival disparities by race

and ethnicity," Cancer, vol. 115, no. 10, pp. 2214–2221, 2009.
[10] W. F. Anderson, P. S. Rosenberg, and S. J. Petito, "Breast cancer trends among women in the United States," Journal of the National Cancer Institute, vol. 102, no. 18, pp. 1–10, 2010.
[11] W. A. Knaus et al., "APACHE II: A severity of disease classification system," Critical Care Medicine, vol. 13, no. 10, pp. 818–829, 1985.
[12] J. L. Vincent et al., "The SOFA (Sepsis-related Organ Failure Assessment) score," Intensive Care Medicine, vol. 22, no. 7, pp. 707–710, 1996.
[13] A. V. Chobanian et al., "Seventh report of the Joint National Committee on Prevention, Detection, Evaluation, and Treatment of High Blood Pressure," Journal of the American Medical Association (JAMA), vol. 289, no. 19, pp. 2560–2572, 2003.
[14] A. H. Mokdad et al., "The continuing epidemics of obesity and diabetes in the United States," JAMA, vol. 291, no. 10, pp. 1238–1245, 2004.
[15] J. W. R. Twisk, Applied Longitudinal Data Analysis for Epidemiology: A Practical Guide. Cambridge, UK: Cambridge University Press, 2003.
[16] P. Diggle, P. Heagerty, K. Y. Liang, and S. Zeger, Analysis of Longitudinal Data, 2nd ed. Oxford, UK: Oxford University Press, 2002.
[17] T. M. Therneau and P. M. Grambsch, Modeling Survival Data: Extending the Cox Model. New York, NY, USA: Springer, 2000.
[18] A. Iasonos, J. Schrag, D. Raj, and P. B. Panageas, "How to build and interpret prognostic models for survival outcomes," Journal of Clinical Oncology, vol. 26, no. 8, pp. 1364–1370, 2008.
[19] C. Curtis et al., "The genomic and transcriptomic architecture of breast tumors," Nature, vol. 486, no. 7403, pp. 346–352, 2012.
[20] C. E. DeSantis et al., "Breast cancer statistics, 2013," CA: A Cancer Journal for Clinicians, vol. 64, no. 1, pp. 52–62, 2014.
[21] R. L. Siegel, K. D. Miller, and A. Jemal, "Cancer statistics, 2019," CA: A Cancer Journal for Clinicians, vol. 69, no. 1, pp. 7–34, 2019.
[22] E. W. Steyerberg et al., "Assessing the performance of prediction models," Statistics in Medicine, vol. 29, no. 3, pp. 230–238, 2010.
[23] M. W. Kattan, "Nomograms are superior to staging and risk grouping systems for predicting survival in oncology," Seminars in Oncology, vol. 30, no. 2, pp. 231–238, 2003.
[24] H. Ishwaran, U. B. Kogalur, E. H. Blackstone, and M. S. Lauer, "Random survival forests," Annals of Applied Statistics, vol. 2, no. 3, pp. 841–860, 2008.
[25] C. Lee, W. Zame, J. Yoon, and M. van der Schaar, "Deep learning for survival analysis: A review," Journal of Biomedical Informatics, vol. 78, pp. 103–114, 2018.