

Article

Not peer-reviewed version

A Unified Framework for Survival Prediction: Combining Machine Learning Feature Selection with Traditional Survival Analysis in Heart Failure and METABRIC Breast Cancer

[Fangya Tan](#) , [Jian-Guo Zhou](#) , Shuqiao Li ^{*} , Bowen Long , Srikar Bellur , [Yang Zhou](#) , Mark Newman

Posted Date: 29 January 2026

doi: 10.20944/preprints202601.2325.v1

Keywords: survival analysis; machine learning; feature selection; gradient boost machine; random survival forest; heart failure; METABRIC breast cancer; hazard ratio; kaplan meier curve; clinical trials; Cox model



Preprints.org is a free multidisciplinary platform providing preprint service that is dedicated to making early versions of research outputs permanently available and citable. Preprints posted at Preprints.org appear in Web of Science, Crossref, Google Scholar, Scilit, Europe PMC.

Copyright: This open access article is published under a [Creative Commons CC BY 4.0 license](#), which permit the free download, distribution, and reuse, provided that the author and preprint are cited in any reuse.

Disclaimer/Publisher's Note: The statements, opinions, and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions, or products referred to in the content.

Article

A Unified Framework for Survival Prediction: Combining Machine Learning Feature Selection with Traditional Survival Analysis in Heart Failure and METABRIC Breast Cancer

Fangya Tan ¹, Jian-Guo Zhou ², Shuqiao Li ^{1,*}, Bowen Long ¹, Srikar Bellur ¹, Yang Zhou ³ and Mark Newman ¹

¹ Department of Analytics, Harrisburg University of Science and Technology, Harrisburg, PA, USA

² Department of Oncology, The Second Affiliated Hospital of Zunyi Medical University, Zunyi, China

³ Department of Physics, The Graduate Center, City University of New York, New York, USA

* Correspondence: sli37@my.harrisburg.edu

Abstract

Background: The clinical adoption of machine learning (ML) in survival analysis is often hindered by the "black box" nature of complex algorithms. This study presents a unified and clinically grounded framework that integrates ML-based feature selection with traditional survival analysis to bridge the gap between algorithmic predictive power and routine clinical interpretation. **Methods:** We employed a hybrid approach where ML models identified high-impact features, which were subsequently validated using Cox Proportional Hazards and Kaplan–Meier analysis. The framework was evaluated across two distinct disease domains, Heart Failure and the METABRIC breast cancer cohort, to assess robustness and generalizability. **Results:** In the Heart Failure dataset, risk stratification based on age group, serum creatinine, and blood pressure successfully separates patients into distinct risk groups. The high-risk group exhibited significantly increased mortality compared to the low-risk group (Hazard Ratio [HR]: 2.61; 95% CI: 1.42–4.78; $p = 0.0013$). Validation in the METABRIC dataset confirmed the adaptability of the method; a composite risk profile utilizing age group at diagnosis, HER2 status, and the Nottingham Prognostic Index (NPI) yielded strong separation ($p < 0.001$). The high-risk breast cancer group demonstrated an HR of 2.73 (95% CI: 2.34–3.19), with a median survival of 104.7 months compared to 252.3 months for the low-risk group, a survival benefit of approximately 12.3 years. **Conclusions:** This cross-disease validation demonstrates that integrating ML feature selection with established survival metrics translates complex models into concise, statistically robust, and clinically interpretable prognostic factors, offering a scalable methodology for diverse clinical contexts.

Keywords: survival analysis; machine learning; feature selection; gradient boost machine; random survival forest; heart failure; METABRIC breast cancer; hazard ratio; kaplan meier curve; clinical trials; Cox model

1. Introduction

Survival analysis is a foundational tool in clinical research, used to model time-to-event outcomes such as disease progression, treatment response, and patient mortality[1,2]. Traditionally, it has been most widely applied in oncology [3,4,5], where time-to-death metrics are critical for evaluating treatment efficacy. More recently, survival analysis has expanded into other therapeutic areas, including cardiovascular disease [6,7,8], liver cirrhosis [9,10]11], and chronic conditions where long-term outcomes or organ failure risks are key concerns. Common methods such as the hazard ratio (HR) and Kaplan–Meier (KM) curves are central to these analyses, enabling clear comparisons

between patient subgroups and supporting both regulatory decisions and patient treatment strategies based on risk [2,12,13].

In parallel with classic statistical approaches like HR and KM, the rise of electronic health records and high-dimensional data has driven rapid adoption of machine learning (ML) in healthcare. ML techniques, ranging from Cox proportional hazards extensions [14] to Random Survival Forests (RSF) [15], Gradient Boosting Machines (GBM)[16], XGBoost [17], and CatBoost [18], have been increasingly used to monitor chronic diseases, identify early biomarkers [19], and predict clinical outcomes such as adverse events [20] and overall survival [21]. These models excel at handling high-dimensional and longitudinal data, making them well-suited for complex prediction tasks. However, their “black-box” nature often limits clinical interpretability, particularly when trying to link model-identified features with actionable, patient-level disease guidance.

By contrast, traditional statistical methods such as the hazard ratio offer clear, interpretable risk comparisons between groups (e.g., treatment vs. placebo) at any time point [5]. Kaplan–Meier curves further illustrate survival probabilities over time and allow for intuitive visualization of group differences [22]. For example, a study of 3,878 breast cancer patients treated in Edinburgh over a 20-year period used HRs and KM curves to show that median survival declined from 10 years (Stage I) to less than 1 year (Stage IV), with mortality differences by stage diminishing after 5 years. Another study demonstrated improved survival following breast cancer recurrence from 1974 to 2000 and identified Estrogen Receptor (ER) status as a key prognostic factor, such that ER-negative patients had a 71% higher risk of death (HR = 1.71) than ER-positive patients ($p < 0.001$) [23].

Similarly, in non-small cell lung cancer, HRs showed that low serum albumin (<35 g/L) was associated with a 42% reduction in risk (HR = 0.588), while $\geq 10\%$ weight loss increased mortality risk by 62% (HR = 1.62) [24]. These examples highlight the enduring value of traditional survival analysis: it provides quantifiable, interpretable risk estimates supported by visual summaries like HR and KM curves.

This principle of using interpretable metrics is foundational even in large-scale evidence synthesis. For example, the systematic review by a comprehensive study [9], analyzing 118 studies, used aggregated evidence from HR-based models to establish key prognostic factors in cirrhosis, such as serum creatinine, total bilirubin, and INR. A key question is whether machine learning models trained on raw clinical data would also identify these well-established predictors or instead focus on less interpretable features.

As the volume of available clinical variables grows, including labs, vitals, and comorbidities, it becomes increasingly labor-intensive to manually test each predictor using traditional regression. Biostatisticians have turned to variable selection techniques such as LASSO [19,25] to address this issue. Yet, LASSO has notable limitations in capturing the complex structure in high-dimensional clinical data. Because its penalty treats predictors independently and does not explicitly account for correlation, LASSO tends to arbitrarily select one variable from groups of correlated features while excluding others. This behavior can lead to unstable feature selection, reduced reproducibility, and the omission of clinically meaningful predictors that convey complementary information [26].

In recent years, ML has gained prominence in clinical research due to its ability to handle high-dimensional data and uncover complex patterns. Its computational power has encouraged researchers to apply it in identifying prognostic features that could guide treatment decisions. For example, study [20] demonstrated that biomarker categories could predict immune-related adverse events (irAEs), offering a way to stratify patients who are more likely to benefit from immunotherapy in non-small cell lung cancer. However, while this study effectively identified predictive biomarkers, it did not explore how irAEs or those biomarkers relate to overall survival outcomes.

Similarly, the EURAMOS-1 [21] study evaluated a rare cancer (osteosarcoma) using eight clinical variables to compare traditional Cox models with Random Survival Forests (RSF) and Survival Neural Networks (SNN), focusing on predictive performance via the concordance index (C-index). In this relatively low-dimensional setting (~2,000 patients, 8 variables), Cox models performed

comparably to more complex machine learning methods. However, the study did not report hazard ratios or offered insights into the clinical significance of the selected features.

These studies collectively highlight a key limitation in many ML applications in medicine: uncertainty around the clinical interpretability of the features these models identify. While ML tools such as SHAP values [7] can rank variables by their predictive influence, it is often unclear whether these features also exhibit statistically consistent and interpretable effects, such as those reflected by HR, and the KM curve in traditional survival analysis. Therefore, the degree to which ML-derived predictors align with established clinical risk factors remains a critical but unresolved question.

Some studies have attempted to bridge this gap. For example, in cardiovascular disease research, machine learning models have successfully highlighted serum creatinine and ejection fraction as key predictors of survival outcomes [6]. Although these variables are clinically meaningful, the analyses are typically limited to a single disease area, raising concerns about their applicability across broader, more diverse therapeutic areas. Other research [8] integrated ML with conventional methods, such as logistic regression, Cox models, naïve Bayes, support vector machines (SVM), and artificial neural networks (ANN), to explore risk stratification. Yet these studies often stopped short of examining survival time explicitly or providing interpretable outputs, which limits the usefulness of such methods for clinical decision-making.

Further work [27] expanded the input space to include coexisting medical conditions like hypertension and diabetes, alongside laboratory variables. While this approach improved predictive performance, raising the AUC by 11% across Cox and RSF, it primarily focused on model accuracy on overall survival event rate rather than detailing the clinical relevance of individual predictors. Even large-scale studies, such as one involving over 468,000 patients [25], which compared models including RF, GBM, XGBoost, CatBoost, SVM, LASSO, and extreme learning machines (ELM), emphasized overall survival model prediction performance. Although this study correctly identified ejection fraction, serum creatinine, and blood urea nitrogen (BUN) as top features and effectively stratified patients by risk, it did not present KM plots or HRs to support the clinical impact of those variables.

The gap in machine learning models highlights predictive features but often lacks clinical evidence to explain their impact, as is also evident in breast cancer research using the METABRIC dataset. For example, study [28] applied Cox proportional hazards (Cox PH), RSF, and conditional inference forests (Cforest), the models consistently identified age, tumor size, and number of positive lymph nodes as important prognostic features. Although these variables appeared influential in the models, the study did not provide supporting evidence, such as HR to quantify their survival impact.

In a similar vein, study [14] employed Cox PH, RSF, and DeepHit to evaluate survival outcomes. Classical factors such as estrogen/progesterone receptor (ER/PR) status, tumor stage, and age remained significant in Cox regression, with hormone receptor positivity associated with lower mortality risk. While the study reported p-values for these features, it did not include explicit hazard ratios, limiting the ability to assess the magnitude of effect. Although the findings aligned with clinical expectations and model outputs, they lacked numerical validation through traditional survival analysis methods.

The study [30] took a more model-diverse approach, using GBM, RSF, SVM, and ANN to predict five-year survival, reporting an overall AUC of 0.67. It emphasized clinically relevant variables such as the Nottingham Prognostic Index (NPI), tumor size, stage, ER/PR/HER2 status, and surgery type. In addition, study [31] demonstrated that explainable machine learning models (XGBoost with SHAP values) can outperform traditional Cox regression in predicting breast cancer survival, achieving higher C-index scores. The study effectively highlighted the importance of clinical features such as tumor stage, age, hormone receptor status, HER2, and treatment information through SHAP-based interpretability. However, these studies did not show how these features were linked to survival using KM curves or HRs, which made it difficult to interpret their individual contributions beyond model accuracy.

A more integrative analysis was presented in [32], which compared Cox PH, RSF, gradient-boosted survival (GBS), and survival support vector machines (SSVM). In this study, ML models, particularly RSF, outperformed Cox PH in predictive performance, especially when combining clinical and transcriptomic features. Using SHAP values for model interpretation, the study identified both established clinical factors (e.g., age) and novel genomic features (e.g., ERAS, SLC14A1, LCN15) as top predictors. While the model offered strong interpretability and biological insight, it still did not report hazard ratios or survival curves to validate these findings in a clinically meaningful way.

Across these studies, a consistent limitation pattern emerges while machine learning methods are effective at identifying predictive features and achieving high accuracy in predicting overall survival status, they often lack clinical grounding through traditional survival metrics. The absence of HRs and KM curves makes it difficult to assess whether model-selected features are statistically robust, clinically actionable, and suitable for translation into practice. This disconnect is a significant barrier to adoption, as most ML studies improve prediction within a single disease domain but fall short of validating their findings with clinically trusted metrics or demonstrating cross-disease applicability. Bridging this methodological gap is essential for advancing the clinical utility of machine learning in survival analysis.

To address the challenges of aligning machine learning outputs with clinically interpretable survival metrics, the present study integrates ML-based feature selection with traditional survival analysis across multiple therapeutic areas. We focus on two chronic diseases, heart failure and breast cancer. The diseases are chosen for their high clinical burden, biological heterogeneity, and the potential for improved outcomes through early risk stratification. Each dataset was selected based on its high quality, transparency, and public availability, ensuring reproducibility and external validation.

- Heart failure remains a leading cause of cardiovascular mortality globally, accounting for an estimated 17.9 million deaths annually [33]. Early identification of high-risk patients is critical for guiding treatment and reducing hospital readmissions. We utilize a dataset comprising medical records of 299 patients collected from the Faisalabad Institute of Cardiology and Allied Hospital in Pakistan between April and December 2015 [34].

- Breast cancer is the most frequently diagnosed cancer among women and the second leading cause of cancer-related death worldwide [22]. Despite advances in screening and treatment, survival disparities persist, underscoring the need for personalized, risk-based care. The METABRIC dataset offers a comprehensive molecular and clinical profile of nearly 2,509 breast cancer patients, combining genomic, transcriptomic, and survival data. Its use has enabled the identification of 10 novel molecular subtypes, significantly enhancing our understanding of prognosis and disease heterogeneity [35].

This study proposes a grounded framework that bridges the gap between machine learning predictions and traditional survival analysis. By validating ML-identified features with standard survival metrics, specifically hazard ratios and Kaplan–Meier curves, we enable clinicians to distill complex algorithms into a small set of statistically robust and clinically interpretable high-impact risk factors. This approach facilitates a shift from reactive, generalized care toward proactive, targeted intervention. Furthermore, we address critical translational challenges, including EHR integration and alert fatigue, ensuring these models function not just as theoretical benchmarks, but as practical, trustworthy tools within existing clinical workflows.

2. Materials and Methods

2.1. Study Dataset, Population, and Data Preprocessing

This study utilized two publicly available datasets: a Heart Failure clinical records cohort comprising 299 patients with 10 features [34] and the METABRIC breast cancer cohort, including 1,310 patients with 33 features after data cleaning [9]. The Heart Failure dataset represents a complete clinical record set with no missing value. To maintain methodological consistency and ensure data

integrity across both datasets, a complete case analysis (CCA) strategy was adopted for the METABRIC dataset. Records containing missing values were excluded to ensure that the benchmarking of machine learning algorithms was conducted on strictly ground-truth clinical observations, avoiding the introduction of synthetic noise or uncertainty associated with data imputation techniques.

The primary outcome for both datasets was overall survival, defined as the time from diagnosis to death or censoring at the last date of follow-up. Non-informative identifiers, including subject ID and cohort indicators, were excluded from the analysis. Variables with no variability or limited informational value were also removed; specifically, cancer type and sex were excluded from the METABRIC breast cancer dataset due to their uniform distribution within the cohort. To reduce redundancy and minimize potential bias in model estimation, highly correlated variables, such as relapse-free survival status and relapse-free survival time, were excluded from the METABRIC dataset. Correlation patterns among candidate variables are presented in Supplementary Figure S1.

2.2. Survival Prediction Modeling

We employed survival modeling approaches specifically designed for right-censored time-to-event data. Unlike standard prediction models that rely on a single target variable, outcomes were jointly defined by overall survival time and a binary event indicator (death vs. censoring). For each dataset, observations were partitioned into an 80% training set for model development and a 20% testing set for independent evaluation. The split was performed using the `sample.split` function to stratify by event status, ensuring that the proportion of events (death vs. censoring) was balanced consistently across the training and testing partitions. However, beyond this outcome-based stratification, no artificial resampling or balancing of clinical features was applied. The natural distribution of patient covariates was strictly preserved to maintain the clinical ground truth. Fixed random seeds were used to ensure reproducibility (Heart Failure: seed = 137; METABRIC: seed = 18).

Three machine learning approaches with established utility in survival analysis were evaluated using R statistical software (version 2025.09.2). Tree-based ensemble models, including Random Survival Forests (RSF) and Gradient Boosting Machines (GBM), were employed to capture non-linear effects and complex interactions. RSF extends random forests to time-to-event data through ensemble hazard estimation [15], while GBM optimizes the Cox partial likelihood to model survival relationships [36]. Additionally, LASSO-penalized Cox regression was included as a regularization-based extension of traditional survival analysis, enabling simultaneous feature selection and prediction via coefficient shrinkage [37]. RSF, GBM, and LASSO models were implemented using the `randomForestSRC`, `gbm`, and `glmnet` packages, respectively. These models were selected to balance non-linear flexibility with clinical interpretability; consequently, Support Vector Machines were excluded due to limited transparency. Hyperparameters were optimized using five-fold cross-validation on the training set, with Kaplan–Meier estimation and log-rank testing (via the `survival` and `survminer` packages) used for downstream clinical validation.

2.3. Model Performance Evaluation

To assess model performance, we employed two complementary and widely accepted metrics for survival analysis. Harrell's concordance index (C-index) was used to evaluate the global ability of each model to correctly rank patient survival times across the entire follow-up period, providing a summary measure of overall discrimination [38]. In addition, the time-dependent area under the curve (AUC) was calculated to assess discriminative performance at clinically meaningful time points. For the Heart Failure dataset, an overall time-dependent AUC was evaluated across observed event times. For the METABRIC breast cancer dataset, the overall AUC was selected to align with standard oncological benchmarks and clinical reporting conventions [39].

These two metrics together provide a sufficient and interpretable assessment of model performance, capturing both global ranking ability and time-specific discrimination. Extensive hyperparameter tuning was not pursued, as maximizing predictive accuracy was not the primary

objective of this study. Rather, model performance served a supportive role, ensuring that the models achieved reasonable discrimination while enabling robust feature selection. Importantly, we observed that the relative importance and selection of key features remained stable across models with varying performance levels, indicating that the identified predictors were not driven by marginal gains in discrimination. This approach allowed the study to focus on the clinical relevance and interpretability of machine learning–identified features rather than optimization of predictive metrics alone.

2.4. Feature Importance and Clinical Translation

To translate machine learning outputs into interpretable clinical insights, feature importance scores were extracted from the GBM, RSF, and LASSO models. Top-ranked features that were either consistent across models or highly ranked within specific disease contexts were selected as candidate variables.

To facilitate clinical validation within a conventional survival analysis framework, continuous variables were categorized into clinically meaningful subgroups using established medical standards. This categorization enabled estimation of hazard ratios (HRs) and assessment of the proportional hazards (PH) assumption. The categorization criteria were defined as follows:

Heart Failure Dataset:

- Age at diagnosis: <65 vs. ≥65 years [40].
- Left Ventricular Ejection Fraction (LVEF): Normal (41–75%) vs. Abnormal (<41% or >75%) [41].
- Serum Creatinine: Normal (0.59–1.04 mg/dL for females; 0.74–1.35 mg/dL for males) vs. Abnormal [42].

METABRIC Breast Cancer Dataset:

- Age at diagnosis: <65 vs. ≥65 years [40].
- Nottingham Prognostic Index (NPI): Good (≤ 3.4), Moderate (3.4–5.4), or Poor (> 5.4) [43].
- Positive Lymph Nodes: Low burden (≤ 3) vs. High burden (> 3) [44].
- Tumor Size: Small (< 20 mm) vs. Large (≥ 20 mm) [45].

All available observations were included in Kaplan–Meier and hazard ratio analyses to fully capture observed survival information and preserve clinically meaningful group characteristics. No data balancing or reweighting was performed, as group distributions reflected real-world clinical prevalence, and retaining the original data structure ensured that estimated associations remained representative of observed clinical patterns.

Finally, composite risk groups (low risk vs. high risk) were constructed by combining three clinically significant features identified through this process. Three features were selected for construction of the final low- and high-risk groups to balance prognostic relevance, model stability, and clinical practicality. Limiting the risk model to a small number of consistently identified and routinely available variables improves interpretability and facilitates straightforward application in routine clinical settings. Survival distributions between risk groups were compared using Hazard Ratio, Kaplan–Meier curves, and the log-rank test [3]. The Heart Failure dataset served as the primary analysis, while the METABRIC dataset was used for exploratory validation. The results of these analyses are presented in the following section, highlighting the consistency between machine learning–based feature importance and clinically interpretable survival effects.

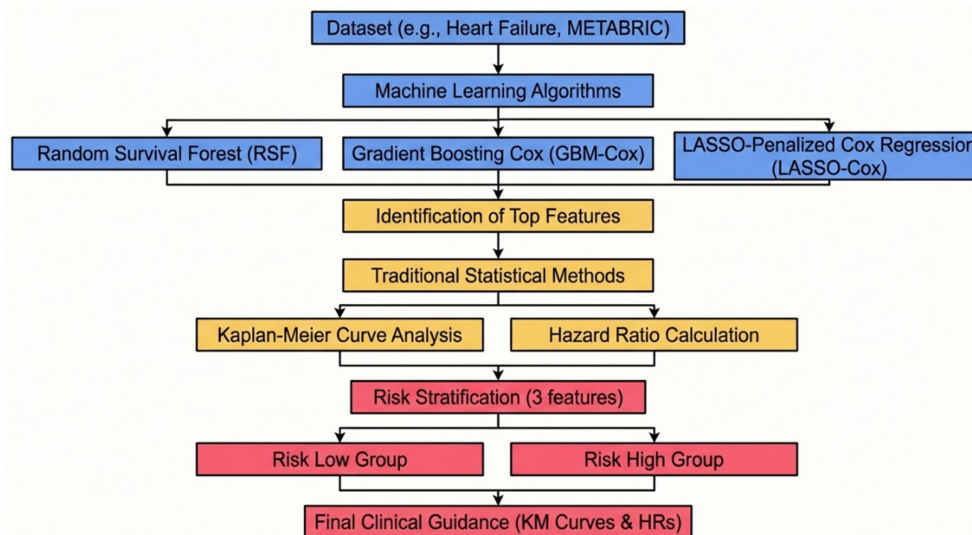


Figure 1. Unified workflow integrating machine learning–based feature selection with survival analysis to enable clinically interpretable risk stratification. The process progresses from machine learning–driven feature selection in high-dimensional data (blue), to identification and statistical validation of key prognostic variables (yellow) and culminates in clinical risk group classification (pink).

3. Result

3.1. Heart Failure Analysis

We evaluated feature importance across three survival modeling approaches: GBM-Cox, RSF, and LASSO-Cox. Across the tree-based models (GBM and RSF), serum creatinine, ejection fraction, and age consistently emerged as the most influential predictors of overall survival time. In contrast, the LASSO-Cox model identified high blood pressure, anemia, and serum creatinine as the primary predictors (Figure 2). The prominence of serum creatinine and ejection fraction as key prognostic factors is consistent with prior studies in heart failure survival analysis [6]. These differences in selected features reflect the underlying algorithmic characteristics of each modeling approach. Tree-based models capture non-linear relationships and interactions among variables, whereas LASSO imposes coefficient shrinkage and favors a sparse linear representation. As a result, LASSO may select high blood pressure as a single representative marker of cardiovascular risk while shrinking correlated variables, such as ejection fraction, toward zero.

Top 10 Feature Importance Across Machine Learning Models for Heart Failure Data

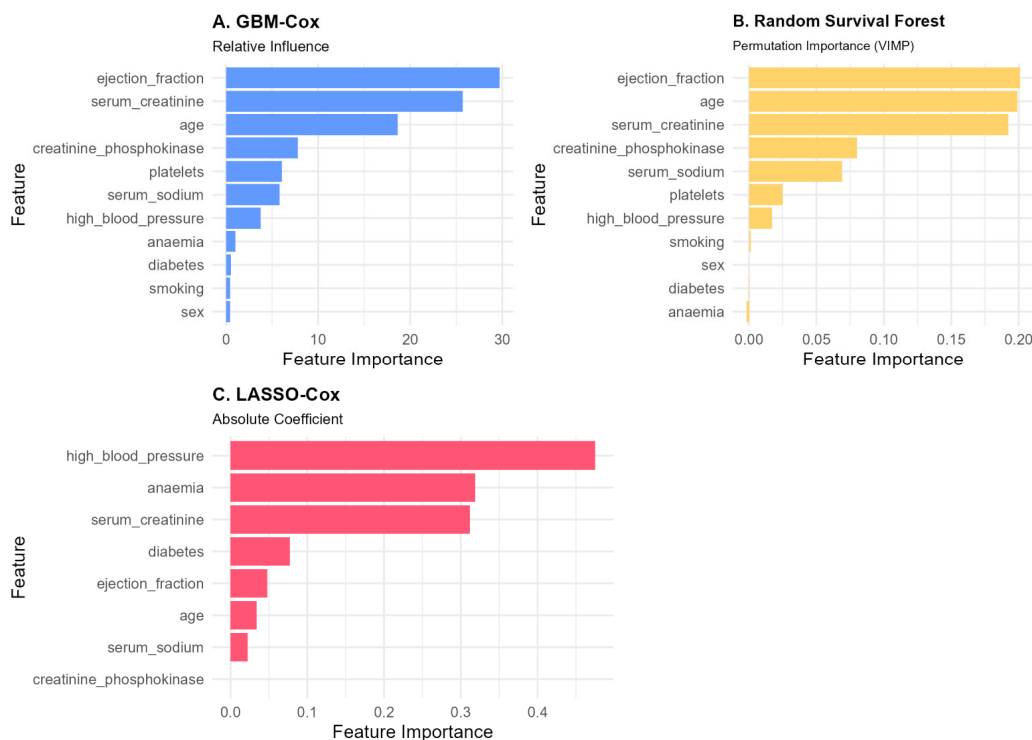


Figure 2. Comparison of Feature Importance across different machine learning models using Heart Failure data. The bar charts illustrate the most significant predictive features identified by (A) GBM-Cox, (B) Random Survival Forest, and (C) LASSO-Cox Model.

Model discrimination performance is summarized in Table 1. In terms of global discrimination, the GBM-Cox model achieved the highest concordance (C-index: 0.789), followed by the Random Survival Forest (RSF) model (0.773) and the LASSO-Cox model (0.731). In contrast, when discrimination was evaluated using time-dependent AUC across the follow-up period, RSF demonstrated the strongest performance (AUC: 0.830), whereas LASSO-Cox (0.724) and GBM-Cox (0.639) showed comparatively lower long-term discrimination. Overall, these estimates fall within the range of C-index values commonly reported in clinical survival studies (approximately 0.72–0.88) [17]. Importantly, the observed level of discrimination indicates that all three models provide sufficient and reliable performance to support downstream analyses. Rather than prioritizing maximal predictive accuracy, the modeling framework was designed to establish stable, interpretable risk factors for the feature-based clinical validation in survival analysis presented in the subsequent section.

Table 1. Discrimination performance of machine learning survival models for Heart Failure Data, summarized using Harrell's concordance index (C-index) and time-dependent area under the curve (AUC).

	GBM-Cox	RFS	LASSO-Cox
Harrell's C-index	0.789	0.773	0.731
All time AUC	0.639	0.830	0.724

Based on the intersection of machine learning-derived feature importance and clinical interpretability, age, serum creatinine, and high blood pressure were selected for downstream

survival validation. Age and serum creatinine were consistently identified as top predictors by both non-linear tree-based models (GBM-Cox and RSF), indicating stable and robust associations with survival time. High blood pressure was additionally selected as a leading predictor in the LASSO-Cox model and represents a routinely monitored clinical variable, supporting the goal of early risk identification and practical patient guidance.

Notably, ejection fraction revealed a divergence between machine learning-based feature importance and traditional regression analysis. Although ranked highly by GBM and RSF, ejection fraction did not reach statistical significance in the univariable Cox model (HR = 1.25, $p = 0.41$). This finding suggests that the effect of ejection fraction on survival may be driven by non-linear or interaction-based mechanisms that are not adequately captured by a linear proportional hazard's framework. Given the study's emphasis on clinical interpretability for the final risk model, ejection fraction was excluded.

Using the clinical categorization criteria defined in the Methods (Section 2.4), Univariable Cox regression (Figure 3) confirmed that all three categorized variables satisfied the proportional hazards assumption and were significant predictors of survival:

- Age (≥ 65 vs. <65 years): $p = 0.0028$ [HR: 1.89];
- Serum Creatinine (Abnormal vs. Normal): $p = 0.0031$ [HR: 1.90];
- High Blood Pressure (Yes vs. No): $p = 0.0189$ [HR: 1.66].

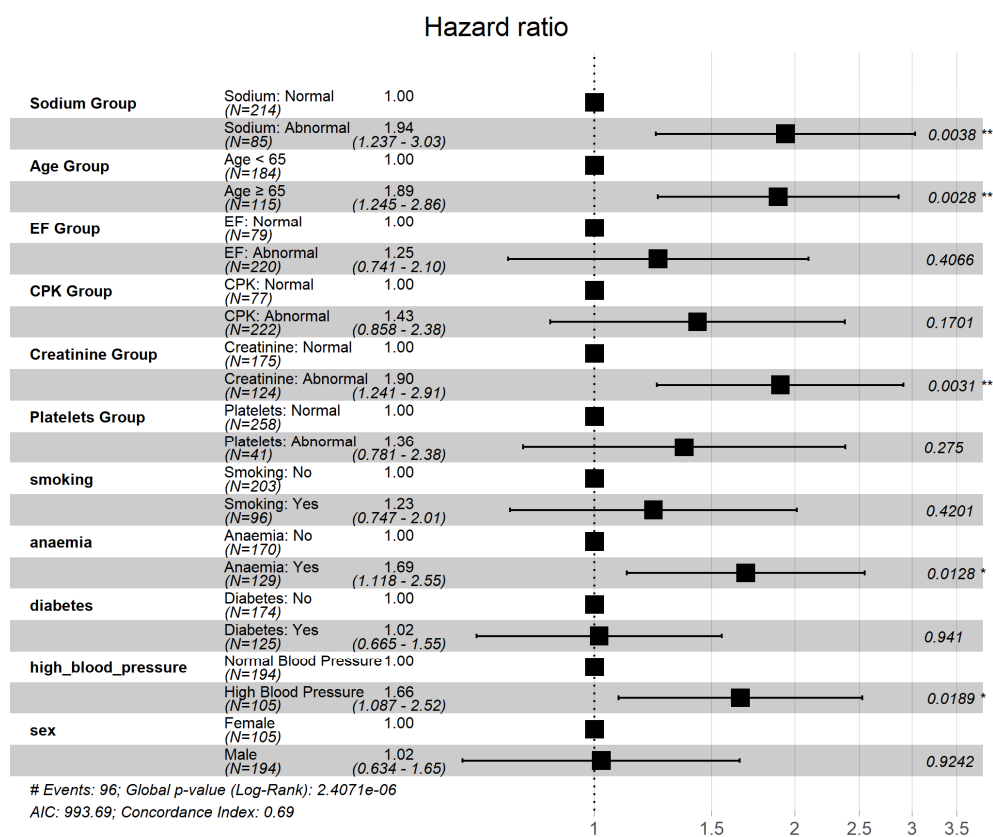


Figure 3. Multivariate Cox Proportional Hazards Model for the Heart Failure Overall Survival. Squares represent hazard ratio estimates and horizontal lines indicate 95% confidence intervals; the vertical dashed line denotes a hazard ratio of 1. p-values are denoted on the right.

Kaplan–Meier survival curves (Figure 4) demonstrated clear separation between risk groups for all selected variables, with statistically significant differences confirmed by log-rank testing (Figure 4). Stratification by age yielded a significant survival difference (log-rank $p < 0.001$); median survival

was 225 days for patients aged ≥ 65 years ($n = 115$), whereas the median was not reached for those < 65 years ($n = 184$). Similarly, patients with abnormal serum creatinine ($n = 124$) exhibited a median survival of 198 days, compared to the normal group ($n = 175$), where the median was not reached (log-rank $p < 0.001$). Survival curves for blood pressure status also showed significant separation (log-rank $p = 0.036$), although median survival was not reached in either the normal ($n = 194$) or high blood pressure ($n = 105$) subgroups.

Collectively, these findings indicate that age, serum creatinine, and blood pressure define clinically distinct risk strata, reinforcing the prognostic relevance of the machine learning-identified features. These individual effects motivated the construction of the composite risk stratification model presented in the following section.

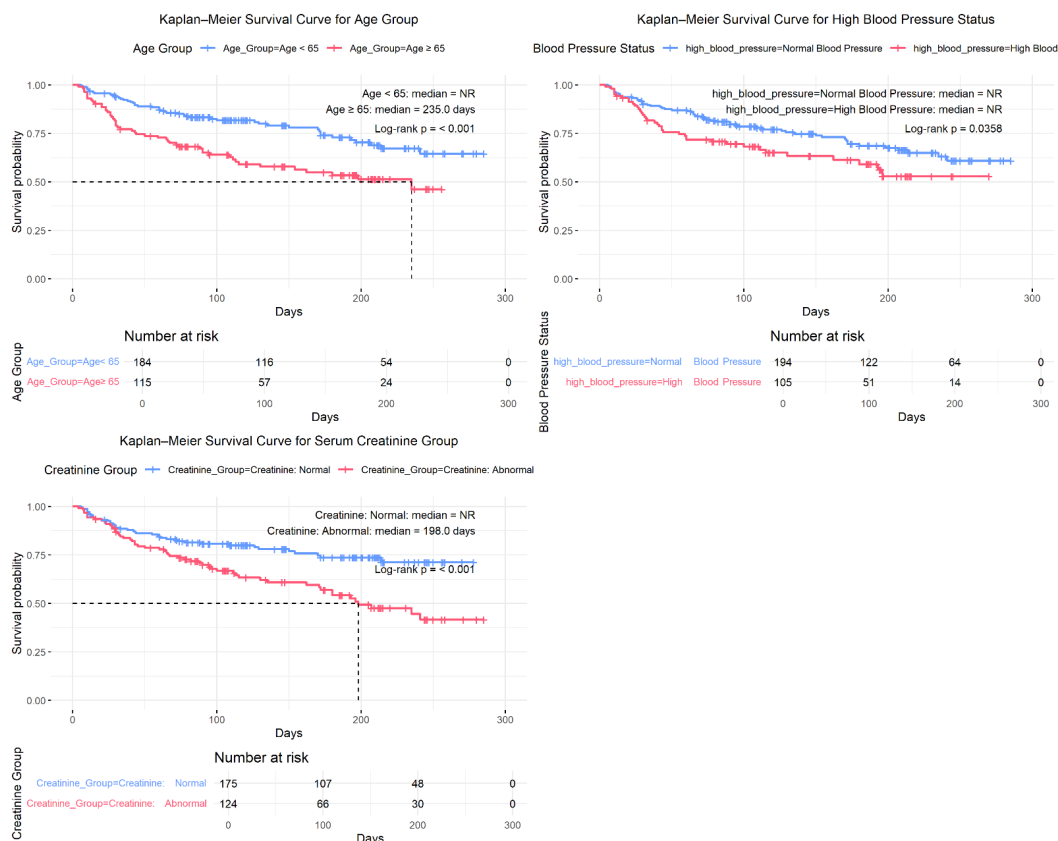


Figure 4. Kaplan-Meier survival curves for overall survival days stratified by constructed Risk group (high vs. low risk) in Heart Failure data. .

The low-risk group consisted of patients aged < 65 years with Normal Serum Creatinine and Normal Blood Pressure; all remaining patients were classified as high risk. Median survival times are shown for each group, and between-group differences were assessed using the log-rank test. The hazard ratio (high vs. low risk) was estimated using a Cox proportional hazards model. Numbers at risk are displayed below the KM curves.

The risk stratification rule was developed using these three variables: age at diagnosis, serum creatinine, and systolic blood pressure. Patients were classified as low risk ($n = 228$) only if all three factors fell within favorable ranges (age < 65 years, with normal blood pressure and serum creatinine as defined in Section 2.4). Conversely, patients presenting with at least one unfavorable factor were classified as high risk ($n = 71$). Kaplan-Meier analysis demonstrated clear separation between the survival trajectories of the two groups (Figure 5), with a statistically significant difference confirmed by the log-rank test ($p = 0.0013$). In Cox regression analysis, the high-risk group exhibited a hazard ratio of 2.61 (95% CI: [1.42 – 4.78]) compared with the low-risk group, indicating a substantially

increased mortality risk. Median survival was not reached in either cohort, reflecting the sustained separation in survival probabilities.

Collectively, these findings demonstrate that combining machine learning–based feature selection with conventional survival analysis yields a simple, interpretable risk stratification tool. By focusing on three routinely available measures, this approach supports the early identification of heart failure patients who may benefit from intensified monitoring and targeted intervention. The study demonstrated how to bridge the gap between data analysis and everyday clinical decision-making.

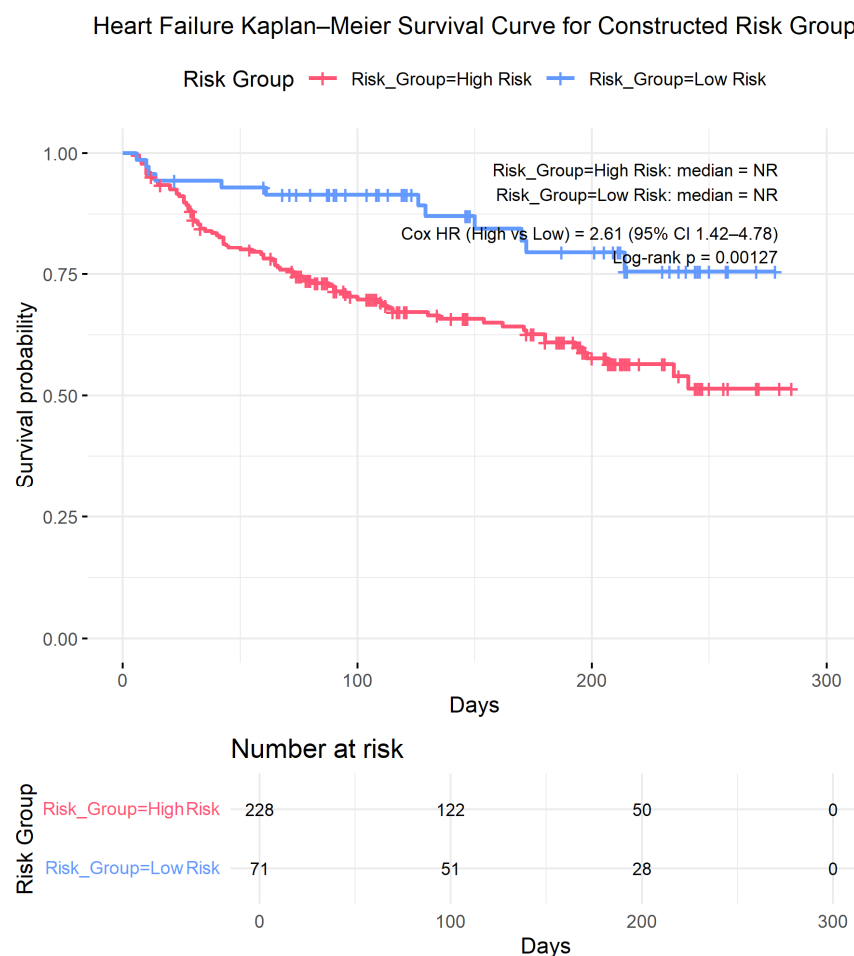


Figure 5. Kaplan–Meier survival curves for overall survival days stratified by constructed Risk group (high vs. low risk) in Heart Failure data. The low-risk group consisted of patients aged <65 years with Normal Serum Creatinine and Normal Blood Pressure; all remaining patients were classified as high risk. Median survival times are shown for each group, and between-group differences were assessed using the log-rank test. The hazard ratio (high vs. low risk) was estimated using a Cox proportional hazards model. Numbers at risk are displayed below the KM curves.

3.2. METABRIC Breast Cancer Analysis

To assess the generalizability of the proposed framework, the same feature selection and survival analysis pipeline was applied to the METABRIC breast cancer cohort as an independent validation dataset. Consistent with the Heart Failure analysis, partial agreement in feature importance was observed across models in the METABRIC breast cancer cohort (Figure 6). The non-linear tree-based models (GBM-Cox and RSF) showed strong concordance for several core prognostic variables, with age at diagnosis, Nottingham Prognostic Index (NPI), tumor size, and number of

positive lymph nodes consistently ranked among the most influential predictors of overall survival time.

In contrast, the LASSO-Cox model emphasized a different subset of variables, prioritizing HER2 status, inferred menopausal state, and Radio Therapy. Notably, NPI emerged as a relatively high-ranked feature in LASSO-Cox models, indicating a shared prognostic signal across modeling paradigms despite differences in feature selection strategies. Importantly, several top-ranked features identified across in tree-based and linear models, including NPI and HER2 status, are well-established and clinically validated prognostic factors in breast cancer [30,31,47].

The observed variation in feature rankings across models likely reflects both algorithmic differences and the increased dimensionality of the METABRIC dataset, which includes a larger number of correlated tumor- and treatment-related variables. Together, these results suggest that while individual models may prioritize different predictors, they converge on a common set of clinically meaningful features, supporting the robustness and generalizability of the proposed framework.

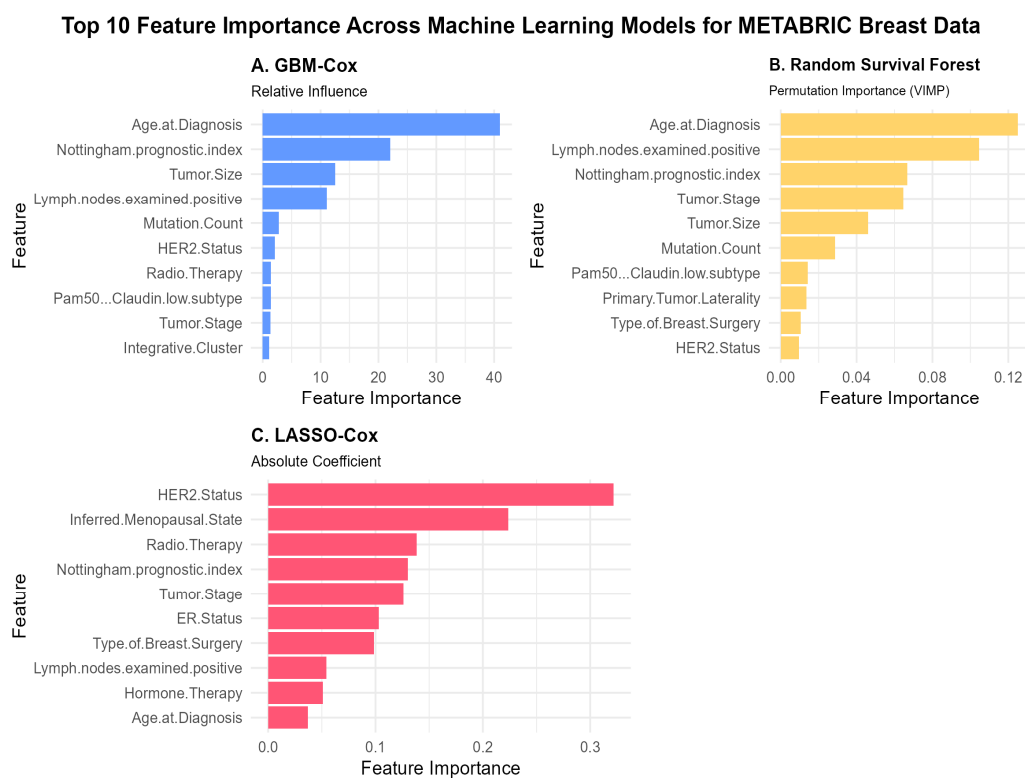


Figure 6. Comparison of Feature Importance across different machine learning models using METABRIC Breast Cancer data. The bar charts illustrate the most significant predictive features identified by (A) GBM-Cox, (B) Random Survival Forest, and (C) LASSO-Cox Model.

Regarding predictive discrimination, tree-based models demonstrated modestly higher performance than the linear approach. RSF and GBM-Cox achieved comparable global ranking accuracy (C-index: 0.681 and 0.679, respectively), slightly exceeding that of the LASSO-Cox model (0.666), with a similar pattern observed for the 5-year time-dependent AUC (0.730 vs. 0.714 vs. 0.675; Table 2).

These performance metrics are comparable to those reported for optimized machine learning survival models, which typically achieve C-index values around 0.7 [30,48]. The model discrimination was used to establish analytical validity rather than as an optimization objective, as the primary focus was the identification of clinically interpretable prognostic features. The modestly

lower discrimination relative to the Heart Failure analysis is expected, given the larger sample size and increased dimensionality of the METABRIC dataset. Accordingly, the prognostic relevance of the selected features was further evaluated using hazard ratios and Kaplan–Meier analyses, as described in the subsequent section.

Table 2. Discrimination performance of machine learning survival models for METABRIC Breast Cancer, summarized using Harrell’s concordance index (C-index) and 5-year time-dependent area under the curve (AUC).

	GBM-Cox	RSF	LASSO-Cox
Harrell’s C-index	0.679	0.681	0.666
AUC (5-year)	0.714	0.730	0.675

We utilize Survival Analysis to validate these findings clinically by assessing the proportional hazards assumption for the top-ranked features [Figure 7]. Inferred menopausal status was excluded due to a lack of statistical significance. Additionally, tumor size and lymph node status were omitted as separate covariates to avoid multicollinearity, as they are intrinsic components of the Nottingham Prognostic Index (NPI) formula [43]:

$$\text{NPI} = (0.2 \times \text{tumor size[cm]}) + \text{lymph node score} + \text{histological grade}$$

Therefore, NPI was retained as a composite measure to capture this prognostic information while maintaining model parsimony. We utilized these three top influential features in ML models to classify patients into risk groups:

- Age at Diagnosis (<65 vs. ≥65 years): $p < 0.001$ [HR: 2.10];
- NPI Group (Good/Moderate vs. Poor): $p < 0.001$ [HR:2.16; HR: 1.49];
- HER2 Status (Negative vs. Positive): $p < 0.001$ [HR: 0.63].

A risk stratification rule was constructed using the three validated prognostic features [Figure 8]. Patients were classified as the Low-Risk group ($n = 621$; age at diagnosis < 65 years, Good/Moderate NPI, and HER2-negative). All remaining patients, presenting with at least one unfavorable risk factor, were classified as the High-Risk group ($n = 689$).

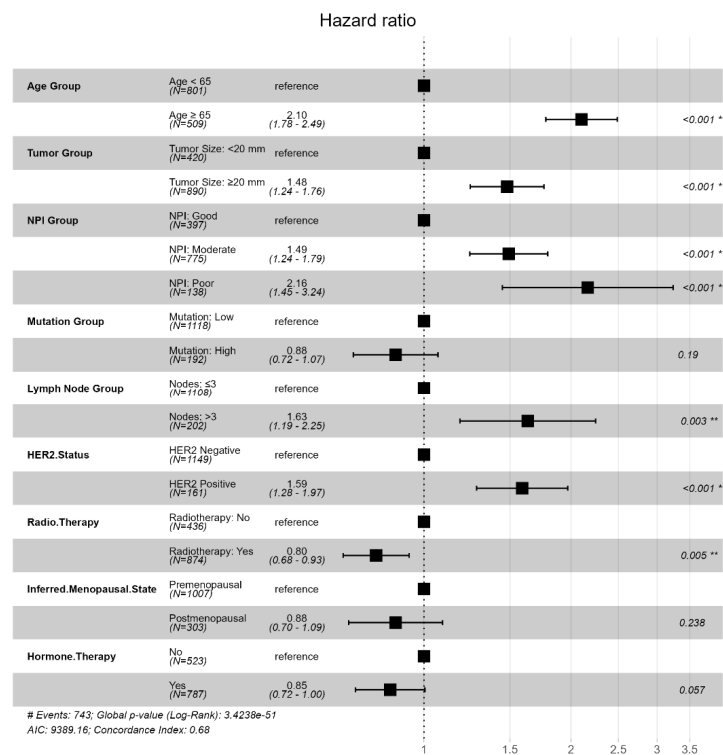


Figure 7. Multivariate Cox Proportional Hazards Model for the METABRIC Breast Cancer Overall Survival. Squares represent hazard ratio estimates and horizontal lines indicate 95% confidence intervals; the vertical dashed line denotes a hazard ratio of 1. p-values are denoted on the right.

Consistent with the Heart Failure, Kaplan–Meier curves in the METABRIC cohort demonstrated clear and statistically significant separation across all three selected variables (Figure 8; log-rank $p < 0.001$ for all). Median survival was substantially shorter among patients' diagnosis age ≥ 65 years compared with those < 65 (109.8 vs. 227.7 months) and among HER2-positive patients compared with HER2-negative patients (96.6 vs. 170.7 months). Stratification by NPI group showed a similar gradient, with median survival ranging from 213.1 months (Good) to 152.1 months (Moderate), and to 59.5 months (Poor).

Kaplan–Meier analysis of the composite risk groups confirmed the robustness of the stratification (log-rank $p < 0.001$) (Figure 9). The high risk group experienced significantly poorer survival (HR=2.73; 95% CI: [2.34 – 3.19]). The median survival difference was pronounced: 104.7 months for the high risk group versus 252.3 months for the low risk group, representing a survival benefit of 147.6 months (≈ 12.3 years).

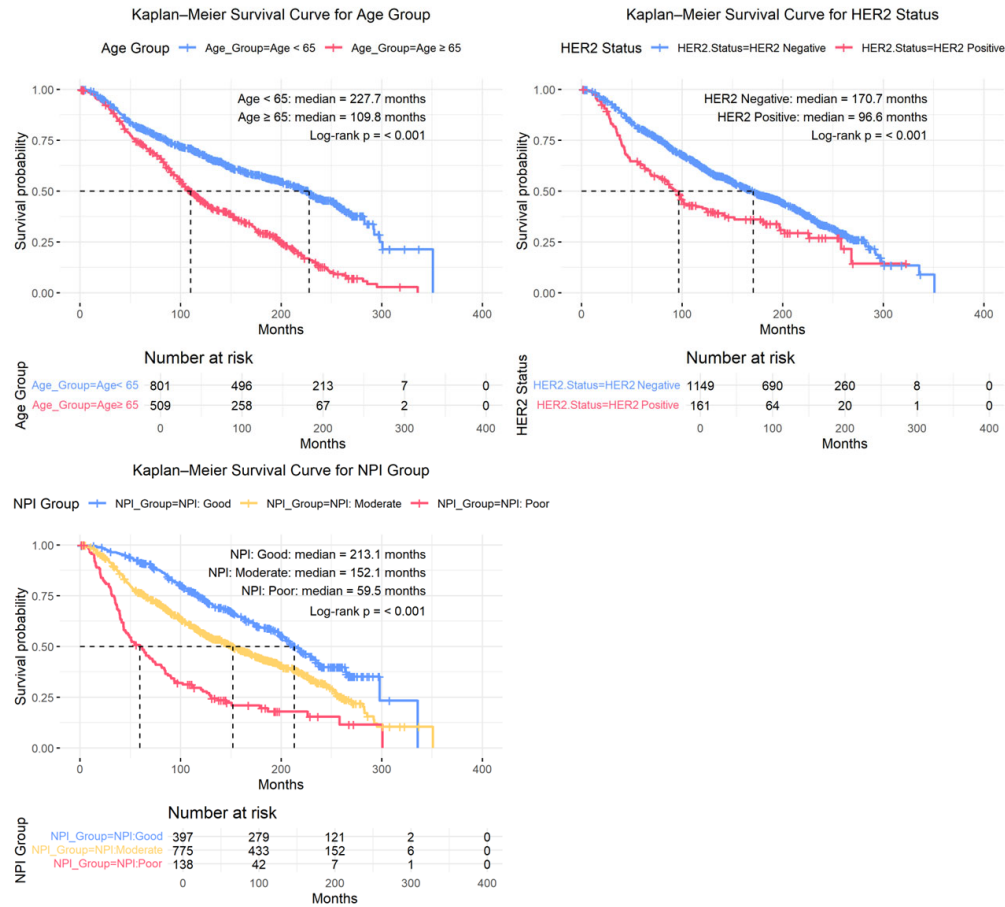


Figure 8. Kaplan-Meier survival curves in METABRIC breast cancer. Survival curves illustrate the probability of overall survival over time (months) stratified by age group at diagnostic, HER2 status, and NPI group. Median survival times are shown within each panel, and between-group differences were assessed using the log-rank test. Risk table including numbers at risk are displayed below each plot.

In summary, the METABRIC analysis validates our methodology in a larger, complex dataset. By integrating machine-learning-derived feature importance with clinical reasoning, the model effectively stratifies breast cancer survival time. Crucially, this supports actionable patient guidance, allowing clinicians to personalize expectations and tailor surveillance strategies based on age at diagnosis, NPI score, and HER2 status.

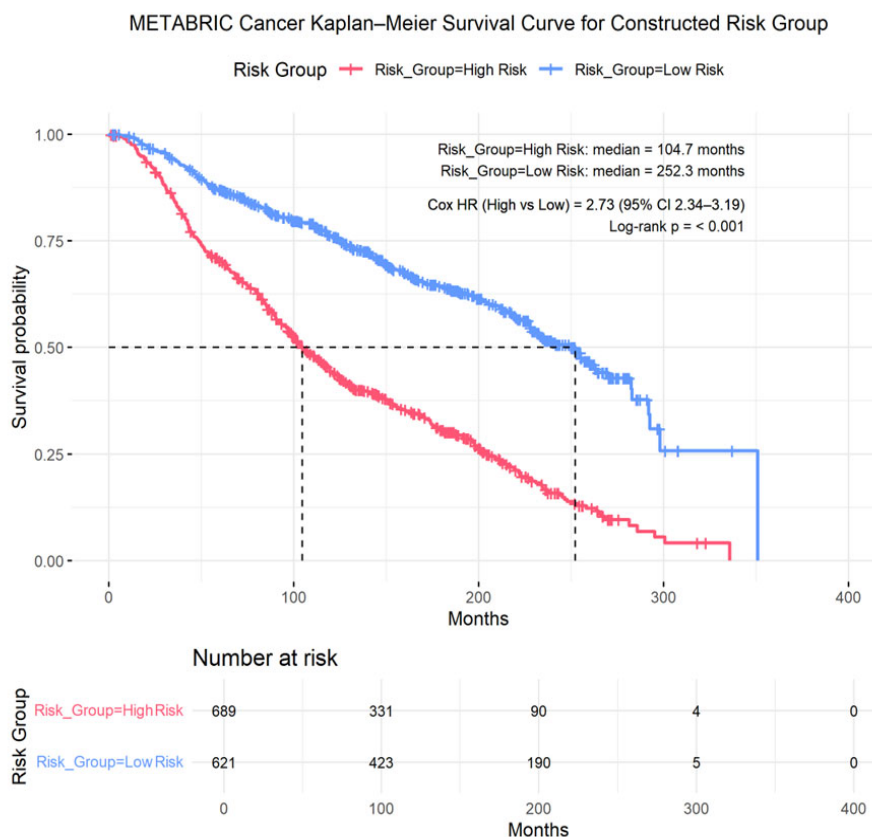


Figure 9. Kaplan–Meier survival curves for overall survival months stratified by constructed Risk group (high risk vs. low risk) in METABRIC Breast Cancer.

The low-risk group consisted of patients aged <65 years at diagnosis with Good or Moderate NPI index and HER2-negative status; all remaining patients were classified as high risk. Median survival times are shown for each group, and between-group differences were assessed using the log-rank test. The hazard ratio (high vs. low risk) was estimated using a Cox proportional hazards model. Numbers at risk are displayed below the KM curves.

4. Discussion

This study demonstrates the utility of integrating machine learning based on feature selection with conventional survival analysis to support clinically interpretable risk stratification in clinical settings. Using heart failure as the primary application and breast cancer as a supportive validation, the proposed framework consistently identified three dominant prognostic features that enabled clear separation of survival risk groups across diseases.

In the heart failure cohort, machine learning models highlighted age, serum creatinine, and blood pressure as the most influential predictors of survival. These variables were subsequently incorporated into a simple risk stratification rule based on routinely collected clinical measurements. Despite the modest sample size and limited feature space, this approach achieved robust discrimination between low- and high-risk patients, illustrating that clinically meaningful prognostic insights can be derived without reliance on complex or opaque model outputs. Importantly, the framework translates abstract machine learning feature importance scores into concrete biomarkers and vital signs that are already embedded in standard clinical workflows, thereby supporting patient monitoring and risk-informed care.

Application of the same analytical framework to the METABRIC [35] breast cancer cohort further demonstrated its adaptability and generalizability. In this larger and more heterogeneous dataset, the framework identified established diagnostic and prognostic factors, including age at diagnosis, Nottingham Prognostic Index (NPI) and HER2 status [31,47]. Risk stratification based on these features again produced clear separation between low- and high-risk groups, reinforcing the disease-agnostic nature of the approach and its potential applicability across diverse clinical contexts. Collectively, these findings support a shift from purely treatment-centric modeling toward comprehensive risk profiling that facilitates earlier identification of patients requiring closer surveillance.

Several limitations warrant consideration. First, the heart failure dataset [34] used in this study contains a relatively small number of observations and a limited set of predictors. Validation using additional heart failure cohorts with larger sample sizes and richer clinical information would further strengthen the robustness of these findings. Nevertheless, validation in the METABRIC dataset, which features substantially greater dimensionality and complexity, provides evidence for the broader applicability of the proposed framework.

Second, although machine learning methods were employed for feature selection, several influential predictors were continuous variables that required categorization to enhance clinical interpretability. While this step may result in some information loss, all cut points were defined a priori based on established clinical guidelines [6] to minimize bias. This highlights the continued importance of clinical expertise in translating data-driven findings into practically useful and actionable tools.

Finally, predictive performance differed across datasets, with stronger discrimination observed in the heart failure cohort than in the METABRIC cohort. This difference likely reflects the greater biological heterogeneity of breast cancer and the complexity of long-term survival outcomes compared to the heart failure cohort. Future work may focus on extending the framework to better accommodate high-dimensional multimodal data while maintaining interpretability, potentially through hybrid or structured modeling approaches.

Despite these limitations, the proposed framework effectively bridges machine learning-based feature discovery with traditional survival analysis. By combining non-linear models, such as Random Survival Forests and GBM-Cox, with penalized regression and validating results using standard Cox regression and Kaplan–Meier analysis, the framework balances diagnostic performance with transparency and clinical relevance.

5. Conclusions

This study presents an interpretable and reproducible framework for connecting machine learning feature selection with conventional survival analysis to support prognostic risk stratification. Across both heart failure and METABRIC breast cancer cohorts, the approach successfully translated machine learning derived important measures into clinically meaningful biomarkers and indicators that can be readily assessed in routine practice.

By prioritizing interpretability and clinical applicability over predictive performance alone, the proposed methodology provides a practical pathway for incorporating machine learning into patient risk profiling and monitoring. This framework supports earlier identification of high-risk patients and more informed clinical decision-making, reinforcing the role of transparent, clinically grounded analytics in advancing precision diagnostics and patient-centered care.

Author Contributions: Conceptualization, F.T. and J.Z.; methodology, F.T. and J.Z. and M.N; software, F.T. and S.L; validation, S.L., Y.Z.; formal analysis, F.T.; investigation, F.T. and S.L and Y.Z.; resources, J.Z.; data curation, F.T.; writing—original draft preparation, F.T.; writing—review and editing, S.L. and B.L; visualization, F.T.; supervision, S.B.; project administration, S.B.; All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable. This study is a secondary analysis of previously published, publicly available, and de-identified datasets and did not involve direct interaction with human participants.

Informed Consent Statement: Not applicable.

Data Availability Statement: The original data presented in the study are openly available in METABRIC breast cancer dataset at [<https://www.cbioportal.org/>] and Heart failure dataset at [ref 34]

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Fleming, T. R., & Harrington, D. P. (1991). *Counting Processes and Survival Analysis*. John Wiley & Sons.
2. Kleinbaum, D. G., & Klein, M. (1996). *Survival Analysis: A Self-Learning Text*. Springer.
3. Kurian, A. W., Sigal, B. M., & Plevritis, S. K. (2010). Survival analysis of cancer risk reduction strategies for BRCA1/2 mutation carriers. *Journal of Clinical Oncology*, 28(2), 222-231.
4. Motzer, R. J., Escudier, B., Tomczak, P., Hutson, T. E., Michaelson, M. D., Negrier, S., ... & Rini, B. I. (2013). Axitinib versus sorafenib as second-line treatment for advanced renal cell carcinoma: overall survival analysis and updated results from a randomised phase 3 trial. *The Lancet Oncology*, 14(6), 552-562.
5. Pocock, S. J., Gore, S. M., & Kerr, G. R. (1982). Long term survival analysis: the curability of breast cancer. *Statistics in Medicine*, 1(2), 93-104.
6. Chicco, D., & Jurman, G. (2020). Machine learning can predict survival of patients with heart failure from serum creatinine and ejection fraction alone. *BMC Medical Informatics and Decision Making*, 20, 1-16.
7. Marcílio, W. E., & Eler, D. M. (2020, November). From explanations to feature selection: assessing SHAP values as feature selection mechanism. In *2020 33rd SIBGRAPI Conference on Graphics, Patterns and Images* (pp. 340-347). IEEE.
8. Mpanya, D., Celik, T., Klug, E., & Ntsinjana, H. (2021). Machine learning and statistical methods for predicting mortality in heart failure. *Heart Failure Reviews*, 26(3), 545-552.
9. DeAmico, G., Garcia-Tsao, G., & Pagliaro, L. (2006). Natural history and prognostic indicators of survival in cirrhosis: a systematic review of 118 studies. *Journal of Hepatology*, 44(1), 217-231.
10. Guo, A., Mazumder, N. R., Ladner, D. P., & Foraker, R. E. (2021). Predicting mortality among patients with liver cirrhosis in electronic health records with machine learning. *PLOS ONE*, 16(8), e0256428.
11. Kanwal, F., Taylor, T. J., Kramer, J. R., Cao, Y., Smith, D., Gifford, A. L., ... & Asch, S. M. (2020). Development, validation, and evaluation of a simple machine learning model to predict cirrhosis mortality. *JAMA Network Open*, 3(11), e2023780.
12. Chen, H. C., Kodell, R. L., Cheng, K. F., & Chen, J. J. (2012). Assessment of performance of survival prediction models for cancer prognosis. *BMC Medical Research Methodology*, 12(1), 102.
13. Spruance, S. L., Reid, J. E., Grace, M., & Samore, M. (2004). Hazard ratio in clinical trials. *Antimicrobial Agents and Chemotherapy*, 48(8), 2787-2792.
14. Lin, D. Y., & Wei, L. J. (1989). The robust inference for the Cox proportional hazards model. *Journal of the American Statistical Association*, 84(408), 1074-1078.
15. Ishwaran, H., Kogalur, U. B., Blackstone, E. H., & Lauer, M. S. (2008). Random survival forests. *The Annals of Applied Statistics*, 2(3), 841-860.
16. Friedman, J. H. (2001). Greedy function approximation: a gradient boosting machine. *Annals of Statistics*, 1189-1232.
17. Chen, T., & Guestrin, C. (2016, August). XGBoost: A scalable tree boosting system. In *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining* (pp. 785-794).
18. Prokhorenkova, L., Gusev, G., Vorobev, A., Dorogush, A. V., & Gulin, A. (2018). CatBoost: unbiased boosting with categorical features. *Advances in Neural Information Processing Systems*, 31.
19. Kokori, E., Patel, R., Olatunji, G., Ukoaka, B. M., Abraham, I. C., Ajekiigbe, V. O., ... & Aderinto, N. (2025). Machine learning in predicting heart failure survival: a review of current models and future prospects. *Heart Failure Reviews*, 30(2), 431-442.

20. Zhou, J. G., Wong, A. H. H., Wang, H., Tan, F., Chen, X., Jin, S. H., ... & Gaipf, U. S. (2022). Elucidation of the application of blood test biomarkers to predict immune-related adverse events in atezolizumab-treated NSCLC patients using machine learning methods. *Frontiers in Immunology*, 13, 862752.
21. Spreafico, M., Hazewinkel, A. D., van de Sande, M. A., Gelderblom, H., & Fiocco, M. (2024). Machine Learning versus Cox Models for Predicting Overall Survival in Patients with Osteosarcoma: A Retrospective Analysis of the EURAMOS-1 Clinical Trial Data. *Cancers*, 16(16), 2880.
22. Watkins, E. J. (2019). Overview of breast cancer. *JAAPA*, 32(10), 13-17.
23. Giordano, S. H., Buzdar, A. U., Smith, T. L., Kau, S. W., Yang, Y., & Hortobagyi, G. N. (2004). Is breast cancer survival improving? Trends in survival for patients with recurrent breast cancer diagnosed from 1974 through 2000. *Cancer*, 100(1), 44-52.
24. Hespanhol, V., Queiroga, H., Magalhaes, A., Santos, A. R., Coelho, M., & Marques, A. (1995). Survival predictors in advanced non-small cell lung cancer. *Lung Cancer*, 13(3), 253-267.
25. Ranstam, J., & Cook, J. A. (2018). LASSO regression. *British Journal of Surgery*, 105(10), 1348.
26. Wang, H., Lengerich, B. J., Aragam, B., & Xing, E. P. (2019). Precision Lasso: accounting for correlations and linear dependencies in high-dimensional genomic data. *Bioinformatics*, 35(7), 1181-1187.
27. Panahiazar, M., Taslimitehrani, V., Pereira, N., & Pathak, J. (2015). Using EHRs and machine learning for heart failure survival analysis. In *MEDINFO 2015: eHealth-enabled Health* (pp. 40-44). IOS Press.
28. Buyrukoğlu, G. (2024). Survival analysis in breast cancer: evaluating ensemble learning techniques for prediction. *PeerJ Computer Science*, 10, e2147.
29. Evangeline I, K., Kirubha, S. A., & Precious, J. G. (2023). Survival analysis of breast cancer patients using machine learning models. *Multimedia Tools and Applications*, 82(20), 30909-30928.
30. Zhao, M., Tang, Y., Kim, H., & Hasegawa, K. (2018). Machine learning with k-means dimensional reduction for predicting survival outcomes in patients with breast cancer. *Cancer Informatics*, 17, 1176935118810215.
31. [31] Moncada-Torres, A., van Maaren, M. C., Hendriks, M. P., Siesling, S., & Geleijnse, G. (2021). Explainable machine learning can outperform Cox regression predictions and provide insights in breast cancer survival. *Scientific Reports*, 11(1), 6968.
32. Doan, L. M. T., Angione, C., & Occhipinti, A. (2022). Machine learning methods for survival analysis with clinical and transcriptomics data of breast cancer. In *Computational Biology and Machine Learning for Metabolic Engineering and Synthetic Biology* (pp. 325-393). Springer US.
33. World Health Organization. (n.d.). World Heart Day. https://www.who.int/cardiovascular_diseases/world-heart-day/en/
34. Ahmad, T., Munir, A., Bhatti, S. H., Aftab, M., & Ali Raza, M. (2017). DATA_MINIMAL [Dataset]. *PLOS ONE*. <https://doi.org/10.1371/journal.pone.0181001.s001>
35. [35] Curtis, C., Shah, S. P., Chin, S. F., Turashvili, G., Rueda, O. M., Dunning, M. J., ... & Caldas, C. (2012). The genomic and transcriptomic architecture of 2,000 breast tumours reveals novel subgroups. *Nature*, 486(7403), 346-352.
36. Vieira, D., Gimenez, G., Marmorela, G., & Estima, V. (2021). XGBoost survival embeddings: Improving statistical properties of XGBoost survival analysis implementation. *Loft Python*.
37. Tibshirani, R. (1997). The lasso method for variable selection in the Cox model. *Statistics in Medicine*, 16(4), 385-395.
38. Pencina, M. J., & D'Agostino, R. B. (2004). Overall C as a measure of discrimination in survival analysis: model specific population value and confidence interval estimation. *Statistics in Medicine*, 23(13), 2109-2123.
39. Kamarudin, A. N., Cox, T., & Kolamunnage-Dona, R. (2017). Time-dependent ROC curve analysis in medical research: current methods and applications. *BMC Medical Research Methodology*, 17(1), 53.
40. U.S. Food and Drug Administration (FDA). (2022). Inclusion of Older Adults in Cancer Clinical Trials: Guidance for Industry. U.S. Department of Health and Human Services.
41. Mayo Clinic. Ejection Fraction: What Does It Mean? Available online: <https://www.mayoclinic.org/tests-procedures/ejection-fraction/about/pac-20384971> (accessed on 24 May 2024).
42. Lab Tests Online. Creatinine. Available online: <https://labtestsonline.org/tests/creatinine> (accessed on 24 May 2024).

43. Galea, M. H., Blamey, R. W., Elston, C. E., & Ellis, I. O. (1992). The Nottingham Prognostic Index in primary breast cancer. *Breast Cancer Research and Treatment*, 22(3), 207–219.
44. Early Breast Cancer Trialists' Collaborative Group (EBCTCG). (2005). Effects of chemotherapy and hormonal therapy for early breast cancer on recurrence and 15-year survival: an overview of the randomised trials. *The Lancet*, 365(9472), 1687-1717.
45. Edge, S. B., Byrd, D. R., Compton, C. C., Fritz, A. G., Greene, F. L., & Trotti, A. (Eds.). (2010). *AJCC Cancer Staging Manual* (7th ed.). Springer.
46. Bland, J. M., & Altman, D. G. (2004). The logrank test. *BMJ*, 328(7447), 1073.
47. Nak, D., & Kivrak, M. (2024). Mastectomy, HER2 receptor positivity, NPI, late stage and luminal B-type tumor as poor prognostic factors in geriatric patients with breast cancer. *Diagnostics*, 15(1), 13.
48. Baidoo, T. G., & Rodrigo, H. (2025). Data-driven survival modeling for breast cancer prognostics: A comparative study with machine learning and traditional survival modeling methods. *PLOS ONE*, 20(4), e0318167.

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.