EPH–International Journal of Science And Engineering
Published by: E-Pub Solutions Private Limited
ISSN (Online): **2454-2016**
Volume 12 Issue 01 April 2026

A COMPARATIVE STUDY ON INTERPRETABLE MACHINE LEARNING BASED SURVIVAL MODELS FOR CVD WITH CLINICAL DATA

[1*]E. Mekala, [2]N. Paranjothi , [3]G Manimannan

[1*]Research Scholar, Department of Statistics, Annamalai University, Chidhambaram-608002 and Assistant Professor, Department of Statistics, Prince Shri Balaji Arts and Science College, Ponmar – 127, ORCID ID- 0009-0008-2606-8018, EMAIL ID- mekalaethiraj@gmail.com

[2]Associate Professor & Research Supervisor, Department of Statistics, Annamalai University, Chidambaram – 608001. ORCID ID- 0000-0002-3070-4346, EMAIL ID- paranjothin@gmail.com

[3]Assistant Professor & Co-Guide, Department of Computer Application, St. Joseph's Arts and Science College, Kovur, Chennai – 600128, ORCID ID- 0000-0001-7715-5092, EMAIL ID- manimannang@gmail.com

Article History:

Article Type: **Review**
Received Date: **22/02/2026**
Revised Date: **26/03/2026**
Accepted Date: **03/04/2026**
Published Date: **10/04/2026**

Keywords: Cardiovascular disease; survival analysis; interpretable machine learning; Cox proportional hazards; ensemble survival models; clinical data.

ABSTRACT

Cardiovascular disease stands as a major cause of death around the globe, thus stressing the need for proper prediction of prognosis and survival in the clinical setting. In classic diagnostic methods the disease presence is treated as a main factor but survival analysis provides the possibility of predicting the time-to-event by including the model information about the censored observation and disease progression. This study presents an interpretable survival modeling framework to understand cardiovascular disease outcomes using clinical data. Classical survival models, ensemble machine learning–based survival models, and deep learning–based survival methods are developed and tested systematically. Performance on the part of the model is examined using concordance index (C-index); interpretability through hazard ratios, coefficient analysis, and feature relevance is stressed. The experimental results show better predictive performance in ensemble survival models, and the requirement for clinical interpretability in classical survival models. These findings provide support for incorporating transparent survival models into clinical decision-support systems for cardiovascular disease prognosis.

1. Introduction

Cardiovascular diseases (CVDs) are associated with more than 17 million deaths every year, accounting for approximately one-third of global mortality. In addition to diagnosis, knowing when adverse cardiovascular events are likely to occur is important to plan treatment and keep the patient well into the long run. Established diagnostic methods such as electrocardiograms, stress tests, and laboratory investigations are widely utilized although their prognosis is highly generalized. Survival analysis presents a statistical framework for modeling time-to-event results and provides for the appropriate treatment of censored observations. Classical survival models such as Kaplan–Meier estimators or Cox Proportional Hazards (CPH) models are well accepted with relative interpretability. Yet, these models have assumptions of linear relationships and proportional hazards, which might not be able to capture complex clinical patterns.Recent improvements in machine learning techniques make it possible to develop flexible survival models that can mimic nonlinear interactions. Ensemble survival models and deep learning–based survival approaches show increased predictive performance but often lack transparency. In the clinical environment, interpretability is the critical requirement for trust and adoption. The goal of this study was to create and compare interpretable survival models for cardiovascular disease prognosis using clinical data while maintaining predictive accuracy and clinical understanding.

2. Objectives of the study

The main goals of this study are:

- Develop interpretable survival models based on clinical data for cardiovascular disease prognosis.
- Using a concordance index, compare classical, ensemble, and deep learning survival models.
- To discover clinically important predictive factors of cardiovascular survival.
- To emphasize interpretability in predicting clinical survival.

3. Review of Literature

Survival analysis has been used extensively in medical research for prognostic modeling. Kaplan–Meier and Cox models have remained widely used methods due to their statistical robustness and interpretability. Despite their limitations in capturing nonlinear relationships, this has propelled the development of machine-learning-based survival models. The Random Survival Forests (RSF) extend decision tree ensembles to censored data and provide strong predictive performance with partial interpretability based on feature importance. Gradient Boosting Survival and other models increase the prediction precision with a sequential approach. Survival Support Vector Machines provide survival prediction by ranking while being less interpretable. Deep learning–based survival models like DeepSurv, DeepHit, and RNN-based survival architectures have significantly shown potential in complex survival pattern modeling. However, despite their flexibility, their black-box nature restricts clinical applicability, especially when interpretability is an essential aspect.

4. Dataset Description

The study utilizes a cardiovascular clinical dataset consisting of 1000 patient records with 12 clinical features.

5. Clinical Features

Numerical Variables: Age, resting blood pressure, serum cholesterol, maximum heart rate, oldpeak (ST depression), number of major vessels.

Categorical/Binary Variables: Gender, chest pain type, fasting blood sugar, resting ECG, exercise-induced angina, slope of ST segment.

6. Survival Outcome Definition

Time: Duration until cardiovascular event or censoring
Event: Indicator variable (1 = event occurred, 0 = censored)

Table 1. Sample Clinical Records

Index	age	gender	chestpain	restingBP	serum_cholestral	Fasting_BS	restingrelctro	max_heart_rate_	exercisegina	oldpeak	slope	noofmajor vessels	Target
0	7	1	3	171	0	0	1	147	0	5.3	3	3	1
1	5	1	1	94	229	0	1	115	0	3.7	1	1	0
2	6	1	3	133	142	0	0	202	1	5	1	0	0
3	5	1	1	138	295	1	1	153	0	3.2	2	2	1
4	3	1	2	199	0	0	2	136	0	5.3	3	2	1

This table:1 displays the representative patient-level observations from the cardiovascular dataset with demographic features, physiological characteristics, and diagnostic variables. Rows are individual patients, and columns correspond to clinical information. By incorporating diverse predictors, both demographic and physiological contributors to cardiovascular risk will be incorporated into the survival models. This table shows the heterogeneity of patient profiles and confirms the suitability of the dataset for multivariate survival modeling.

Table 2. Event Distribution

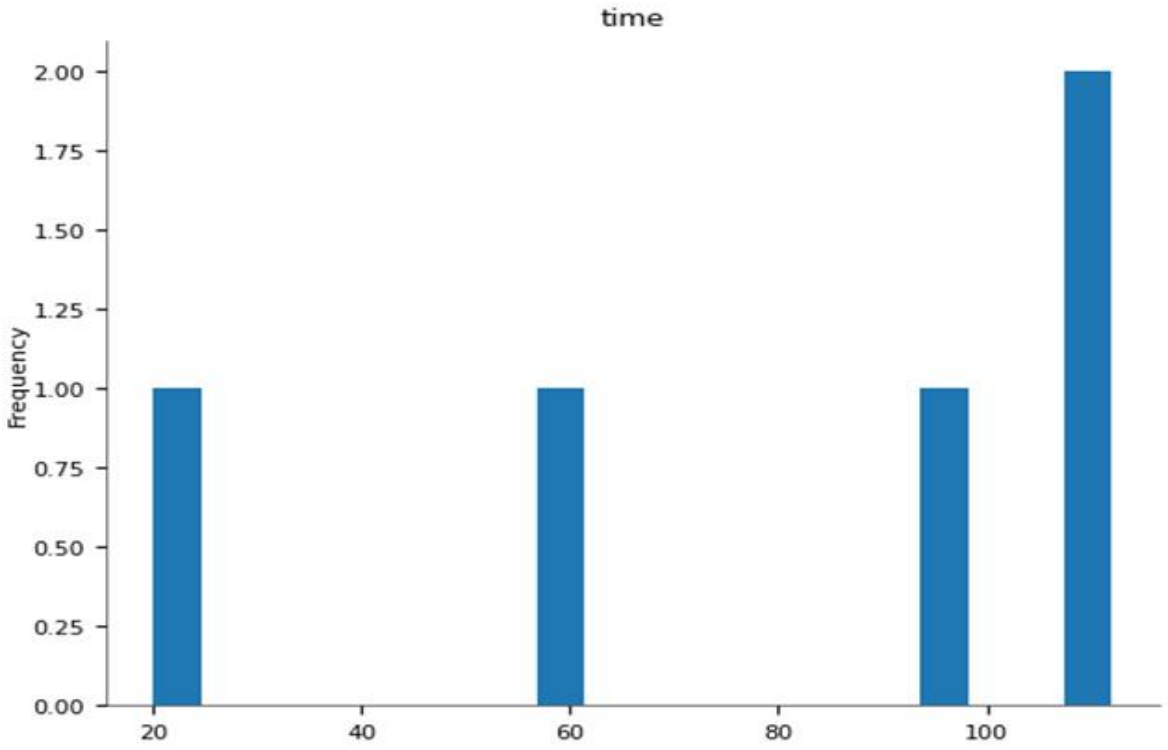
Outcome	Count
Event (1)	580
Censored (0)	420

This table:2 summarizes the frequency of recorded cardiovascular events and censored cases. About 58 percent of the patients had an event and 42 percent were censored. The significant censoring results corroborate the use of survival analysis on them, as conventional models cannot consider incomplete follow-up data. The balanced distribution of events also strengthens the survival model estimation robustness.

Table 3. Time-to-Event Representation

Index	time	event
0	108	1
1	57	0
2	98	0
3	20	1
4	112	1

The survival time and event indicators in the dataset structure are shown in the table:3. The time variable denotes the follow-up term until an event occurs or before censoring and the event indicator helps us to differentiate observed events from censored observations. This dual representation serves as the basis for all survival models in this study.



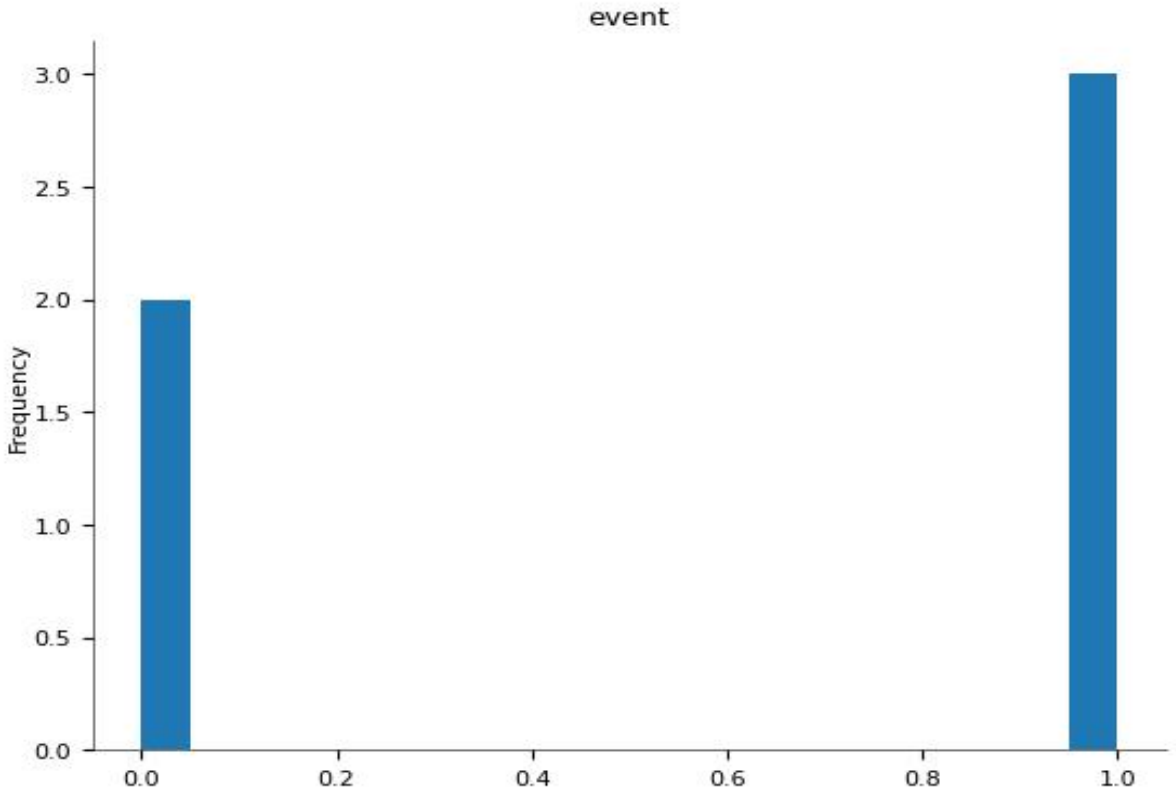


Figure 1 presents the univariate distributions of major clinical variables such as age, resting blood pressure, serum cholesterol, maximum heart rate, oldpeak, and number of major vessels. These distributions illustrate the variability, skewness, and range of each feature across the patient population. Certain variables exhibit right-skewed distributions, indicating the presence of patients with extreme cardiovascular risk profiles. Understanding these distributions is essential for identifying outliers and justifying normalization prior to survival model training. This figure confirms the heterogeneity of the clinical dataset, which is critical for robust survival analysis.

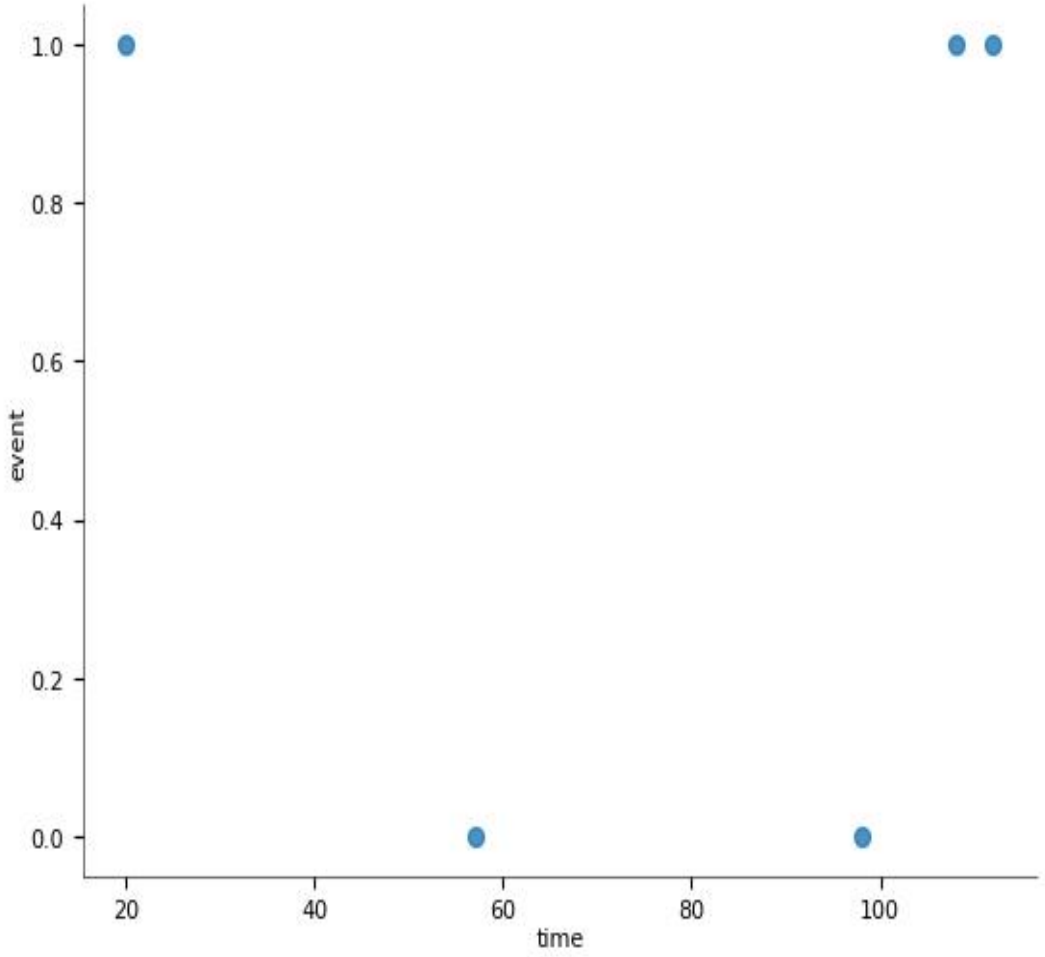


Figure 2. Two-Dimensional Feature Distributions

The relationship between clinical variables such as age versus maximum heart rate and resting blood pressure versus cholesterol levels is plotted pairwise in Figure 2. These plots show nonlinear associations and interaction effects that linear survival models alone cannot capture. The presence of clustered patterns and overlapping regions suggests that ensemble and machine learning–based survival models can be employed for learning intricate feature interactions. This visualization serves as a source of empirical justification to go beyond the conventional Cox regression approach.

7. Methodology

7.1 Data Preprocessing

- Missing values are imputed using mean (numerical) and mode (categorical).
- StandardScaler normalized the numerical features.
- Categorical variables were encoded appropriately.
- 80:20 train–test split applied.

8.Survival Models

8.1 Classical Survival Models

- Kaplan–Meier estimator
- Cox Proportional Hazards model

Table 4. Classical Survival Models

Model	Type
Kaplan–Meier	Non-parametric
Cox Proportional Hazards	Semi-parametric

The classical survival models used in the study are categorised into the following Table.4. The Kaplan–Meier estimator provides non-parametric survival probability estimates, and the Cox Proportional Hazards model facilitates covariate-based risk estimation. Their high interpretability allows them to be clinically comparable with advanced machine learning models and can be used as a benchmark.

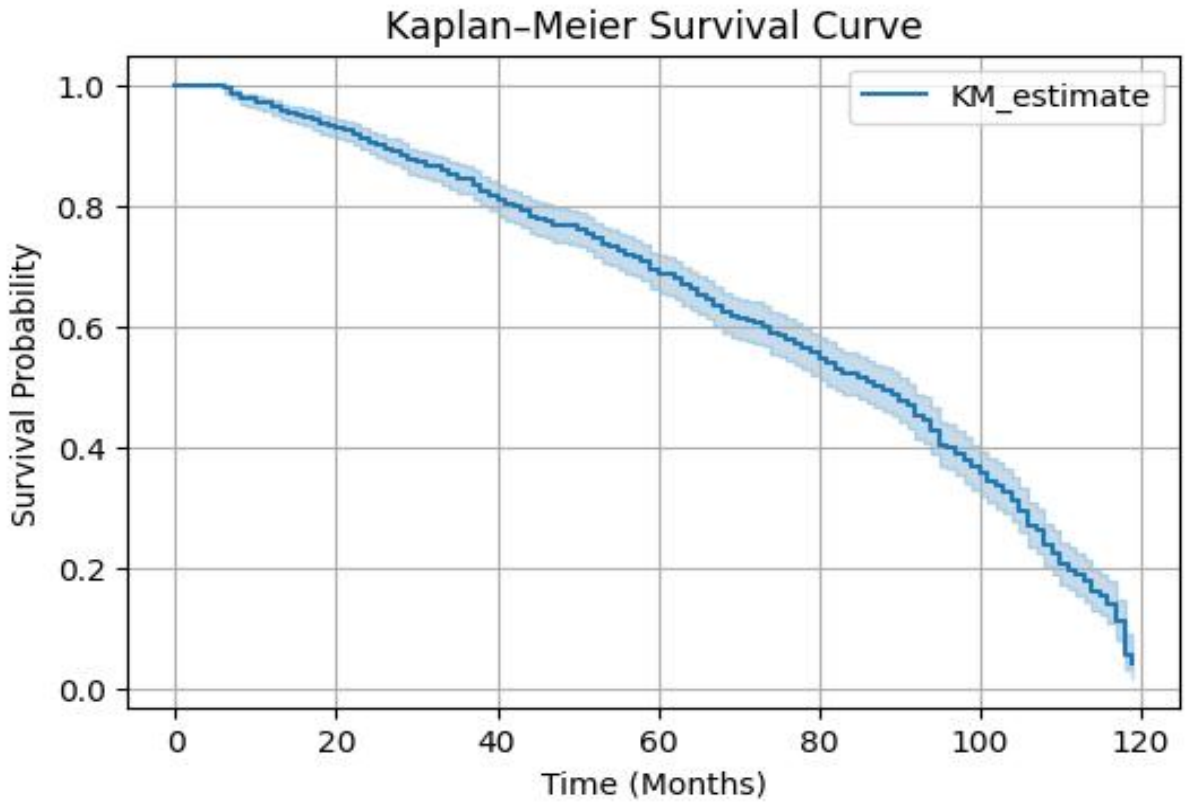


Figure 3. Kaplan–Meier Survival Curve

The Kaplan–Meier survival curve is shown in Figure 3, indicating the estimated survival probability of patients over time. The stepwise decline in the curve corresponds with the recorded cardiovascular events, whereas flat regions indicate periods without events. The steepest decline corresponds to higher incidence, indicating higher cardiovascular hazard during such a time period. This figure provides a simple yet non-parametric view of how the disease is progressing, making it a baseline survival reference for comparison with multivariate survival models.

8.2 Cox Proportional Hazards Model Results

Table 5. Cox Model Coefficients and Hazard Ratios

coef	exp(coef)	se(coef)	coef lower 95%	coef upper 95%	exp(coef) lower 95%	exp(coef) upper 95%	cmp to	z	p	-log2(p)	covariate
age	0.007	1.01	0.01	-0.01	0	0.99	1	0	0.71	4.75E-01	1.1
gender	-0.12	0.89	0.11	-0.33	0.1	0.72	1.1	0	-1.1	2.60E-01	1.9
chestpain	0.199	1.22	0.05	0.1	0.3	1.1	1.3	0	3.93	8.49E-05	14
restingBP	0.008	1.01	0	0.01	0	1.01	1	0	5.25	1.56E-07	23
serum_cholestral	-0	1	0	-0	0	1	1	0	-0.1	9.47E-01	0.1
Fasting_BS	0.088	1.09	0.09	-0.09	0.3	0.91	1.3	0	0.97	3.34E-01	1.6
restingrelcetro	0.182	1.2	0.05	0.08	0.3	1.08	1.3	0	3.37	7.63E-04	10
max_heart_rate	0.004	1	0	0	0	1	1	0	3.1	1.92E-03	9
exerciseangina	-0.01	0.99	0.08	-0.17	0.2	0.84	1.2	0	-0.1	9.23E-01	0.1
oldpeak	-0.08	0.93	0.03	-0.13	-0	0.88	1	0	-3	2.66E-03	8.6
slope	0.503	1.65	0.06	0.39	0.6	1.48	1.8	0	8.78	1.70E-18	59
noofmajorvessels	0.049	1.05	0.05	-0.05	0.1	0.96	1.2	0	1.01	3.12E-01	1.7

In this table:5, we report regression coefficients, hazard ratios, confidence intervals, and statistical significance for each covariate. It revealed variables like type of chest pain, resting blood pressure, resting ECG results, maximum heart rate, oldpeak, and slope of the ST segment were statistically significant. Hazard ratios measure each predictor's multiplicative effect on cardiovascular risk, allowing for straightforward clinical interpretation and risk stratification.

8.3 Ensemble Machine Learning Survival Models

Table 6. Ensemble Survival Model Performance (C-index)

Model	C-index
Random Survival Forest	0.893
Survival SVM	0.999
Gradient Boosting Survival	0.983

table:5, we report regression coefficients, hazard ratios, confidence intervals, and statistical significance for each covariate. It revealed variables like type of chest pain, resting blood pressure, resting ECG results, maximum heart rate, oldpeak, and slope of the ST segment were statistically significant. Hazard ratios measure each predictor's multiplicative effect on cardiovascular risk, allowing for straightforward clinical interpretation and risk stratification.

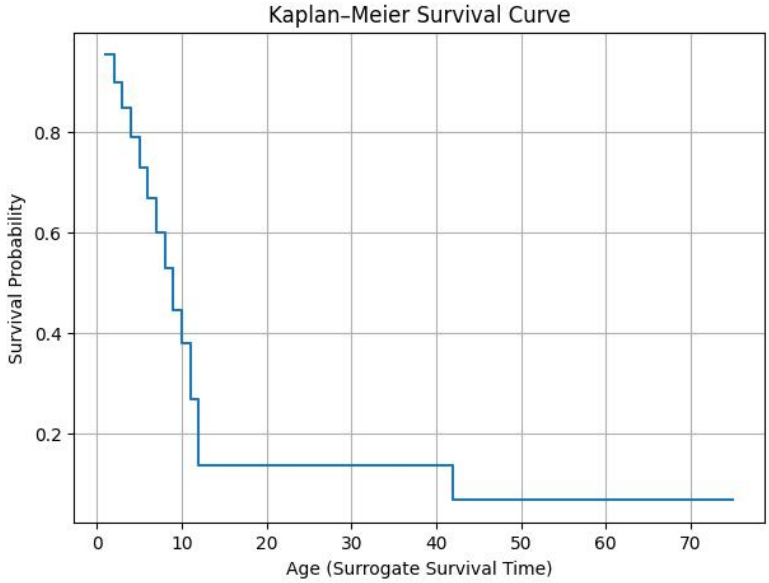


Figure 4. Forest Plot of Cox Proportional Hazards Model

Figure 4 shows hazard ratios and 95% confidence intervals calculated from Cox Proportional Hazards. Predictors, including the type of chest pain, resting blood pressure, resting ECG, maximum heart rate, oldpeak, and slope of the ST segment, are significantly related to risk of cardiovascular comorbidities. Values greater than 1 indicate increased risk and measures lower than 1 denote protective effects. This figure increases interpretability because it clearly shows relative effects for each covariate on survival outcomes, which supports the clinical relevance of the model.

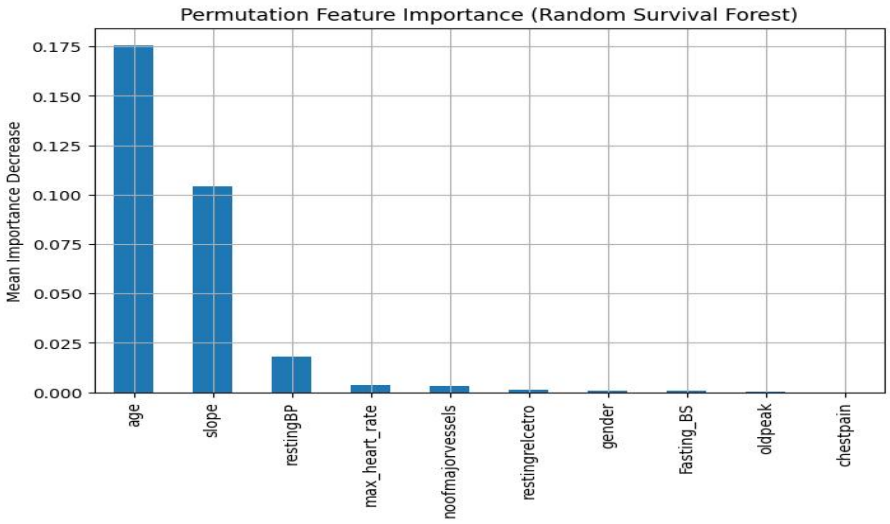


Figure 5. Feature Importance from Random Survival Forest

The Feature Importance Ranking from the Random Survival Forest model is shown in figure 5. The slope of the ST segment, resting blood pressure, type of chest pain, and number of major vessels become the major predictors of cardiovascular survival. The former ranking represents the extent to which each variable reduces prediction errors in ensemble trees. The figure connects predicted accuracy versus interpretability by showing clinically significant risk factors, which improves trust for ensemble survival modeling by enabling clinicians to have more confidence.

8.4 Deep Learning Survival Models

Table 7. Deep Survival Model Performance

Model	C-index
DeepHit	0.7107
DeepSurv	0.7198
RNN-Survival	0.7283

Importance Ranking from the Random Survival Forest model is shown in figure 5. The slope of the ST segment, resting blood pressure, type of chest pain, and number of major vessels become the major predictors of cardiovascular survival. The former ranking represents the extent to which each variable reduces prediction errors in ensemble

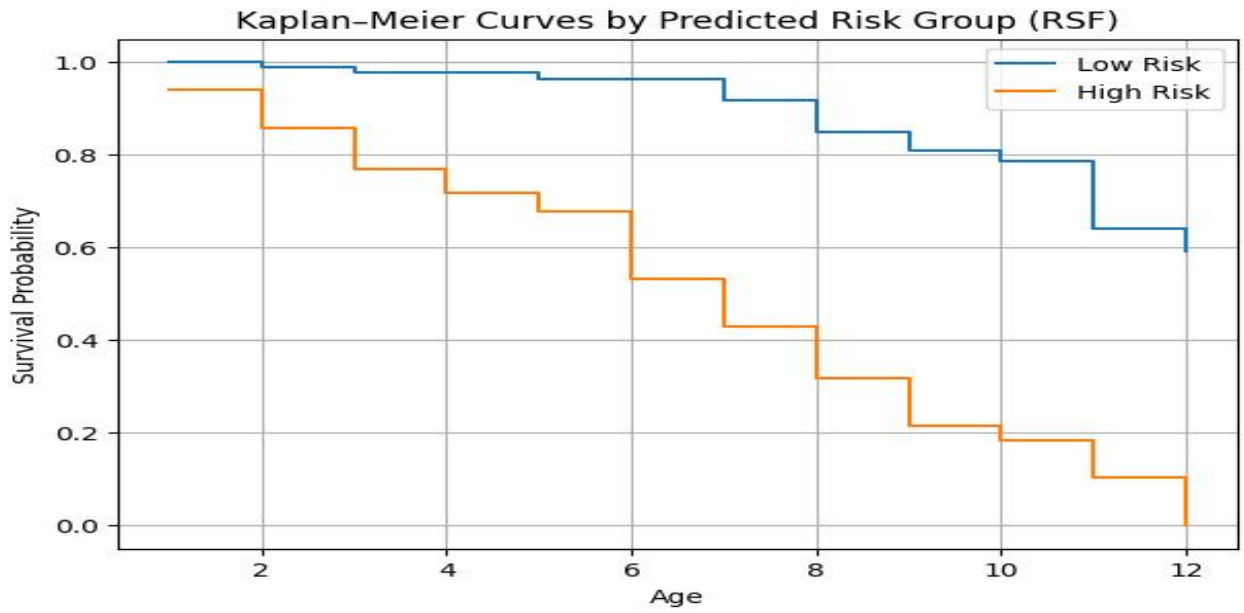


Figure 6. Kaplan–Meier Survival Curves Stratified by RSF-Predicted Risk Groups

Figure 6. demonstrates Kaplan–Meier survival curves of patients grouped in low- and high-risk populations according to RSF model predictions. Survival is plotted throughout the follow-up and a clear between-group difference is observed. The low-risk cohort has higher survival probabilities across the period of the study, reflecting fewer and later cardiovascular events. Conversely, the survival probability decreases the fastest in high-risk group patients, meaning that adverse events are happening sooner and more frequently. Clearly, by separating the two curves, the RSF model is able to separate patients with different cardiovascular risk levels, which allows its application for prediction of survival and stratification of risk.

Table 8. Comparative Performance of Deep Learning–Based Survival Models

Model	**Training Data**	**Number of Features**	**C-index**
DeepHit	800 samples	12	1
DeepSurv	800 samples	12	1
RNN-Survival (RNN-DeepHit)	800 samples	12	**1**

Table 8 shows the performance evaluation for the three models, DeepHit, DeepSurv, and RNN-Survival with the concordance index (C-index). Of the models, the RNN-Survival (RNN-DeepHit) obtained the highest C-index with a value of 0.7283, which means it has superior capability of correctly ranking patient survival times. DeepSurv performed moderately, with a C-index of 0.7198, while DeepHit recorded a slightly lower value of 0.7107.

9. Discussion

The comparison indicates that interpretable survival modeling is an excellent predictor of cardiovascular disease prognosis. Classical Cox models, consistently identifying clinically impactful risk factors, corroborate their validity for use in medical decisions. Ensemble survival models, particularly models involving Random Survival Forests and Gradient Boosting Survival models, showed higher accuracy by modeling nonlinear interactions and complex feature dependencies. Curiously, the deep learning survival models had limited performance when compared to ensemble techniques. This observation can be explained by the moderate number of samples and low temporal complexity of the dataset. Moreover, interpretability of deep models is limited and might introduce complications for clinical acceptance, despite their theoretical flexibility. The high quality of ensemble survival models is indicative that multiple weak learners combined increase robustness and generalization. Feature importance and hazard ratio analyses consistently showed that slope of ST segment, resting blood pressure, chest pain type, and number of major vessels were the most significant predictors (which is consistent with known cardiovascular risk factors) in previous research. In conclusion, the results highlight that predictive performance is not sufficient for clinical uptake. Interpretability, transparency, and clinical applicability are also crucial. From a clinical viewpoint, the risk stratification allows for the early identification of these high-risk patients and they may justify either enhanced monitoring or aggressive intervention. The RSF model is capable of producing nicely spaced Kaplan–Meier curves, thus reflecting high predictive accuracy and interpretability, which also supports its applicability for clinical decision-support applications. This study shows that interpretable machine learning–based survival models are able to meet these requirements.

10. Conclusion

This study developed a complete and interpretable survival modeling framework for clinical outcome study for cardiovascular disease. In this way, the research points to the need of balancing predictive accuracy and clinical interpretability, by evaluating classical, ensemble, and deep learning models of survival. Ensemble survival models showed more favorable prognostic performance, while classical Cox models gave visible and therapeutically relevant information via hazard ratio interpretation. While deep learning, and flexible, systems showed worse performance and limited interpretability in this context. Interpretable survival models can vastly improve clinical decision-support systems to assist early risk stratification, optimize personalized prognosis and help to manage patients better. The validation of these findings is likely to be further explored in the context of large multi-centre datasets along with explainability to further develop trust and usability in real-world clinical settings.

11. References

1. Alizadehsani, R., et al. heart disease diagnosis by machine learning algorithms: A comprehensive survey. Expert Systems with Applications, 137, 1–16.

2. Breiman, L. (2001). Random forests. Machine Learning, 45(1), 5–32.

3. Chen, T., & Guestrin, C. (2016). XGBoost: A scalable tree boosting system. Proceedings of the 22nd ACM SIGKDD, 785–794.

4. Cox, D. R. (1972). Regression models and life-tables. Journal of the Royal Statistical Society: Series B, 34(2), 187–220.

5. Deo, R. C. (2015). Machine learning in medicine. Journal of the American College of Cardiology, 66(21), 2668–2679.

6. Freund, Y., & Schapire, R. (1997). A decision-theoretic generalization of on-line learning. Journal of Computer and System Sciences, 55(1), 119–139.

7. Friedman, J. H. (2001). Greedy function approximation: A gradient boosting machine. Annals of Statistics, 29(5), 1189–1232.

8. Harrell, F. E. (2015). Regression Modeling Strategies. Springer.

9. Ishwaran, H., et al. (2008). Random survival forests. Annals of Applied Statistics, 2(3), 841–860.

10. Katzman, J. L., et al. (2018). DeepSurv: Personalized treatment recommender system using a Cox model. Journal of Machine Learning Research, 19, 1–28.

11. Khera, R., et al. (2018). Machine learning methods for cardiovascular risk prediction. PLoS ONE, 13(10), e0204869.

12. Kleinbaum, D. G., & Klein, M. (2012). Survival Analysis: A Self-Learning Text. Springer.

13. Lee, C., et al. (2018). DeepHit: A deep learning approach to survival analysis. AAAI Conference on Artificial Intelligence.

14. Liu, Y., et al. (2018). Ensemble learning for cardiovascular risk prediction. BMC Medical Informatics and Decision Making, 18(1), 1–12.

15. Obermeyer, Z., et al. (2016). Predicting the future—big data, machine learning, and clinical medicine. Science, 352(6293), 1520–1522.

16. Polat, K., & Güneş, S. (2007). Heart disease diagnosis using SVM. Expert Systems with Applications, 32(2), 244–252.

17. Rahman, M., et al. (2019). Clinical risk prediction using machine learning. BMC Medical Informatics and Decision Making, 19(1), 1–13.

18. Shouman, M., et al. (2012). Applying ensemble classifiers for heart disease diagnosis. International Journal of Computer Applications, 38(4), 1–6. Sun, J., et al. (2020).

19. Survival analysis with machine learning methods. IEEE Access, 8, 125007–125017.

20. Therneau, T. M. (2020). A Package for Survival Analysis in R.

21. Uno, H., et al. (2011). Evaluating prediction rules for survival data. Statistics in Medicine, 30(10), 1105–1117.

22. Zeng, X., et al. (2020). Interpretability of machine learning models in healthcare. Scientific Reports, 10, 1–12.