
COMPARING COX PROPORTIONAL HAZARDS MODEL AND PARAMETRIC MODELS FOR ANALYZING THE SURVIVAL OF PATIENTS WITH HEART FAILURE

Dr. Goutam Barman*

*Department of Statistics,
Krishnagar Government College, West Bengal, INDIA

Email id: goutambrmn@gmail.com

DOI: 10.5958/2249-7137.2025.00040.4

ABSTRACT

Heart failure is still a major global cause of morbidity and death, so assessing patient survival requires strong statistical tools. To investigate how long patients with heart failure live, this study compares the Cox Proportional Hazards (PH) model with common parametric survival models, including Exponential, Weibull, Log-logistic, Log-normal, Gamma, Gompertz, and Rayleigh. We start by checking the proportional hazards assumption of the Cox model using clinical data that covers patient demographics, comorbidities, and survival outcomes. Then, we explore how flexible and efficient the parametric models are for calculating hazard rates and survival functions. We compare the models based on fit metrics such as the Akaike Information Criterion (AIC). While the Cox PH model assumes less about the baseline hazard, our findings show that some parametric models offer better interpretability and predictive accuracy when their assumptions hold true. This comparison highlights the importance of picking the right model for survival studies. Using parametric methods can result in more precise risk assessment for heart failure prognosis.

KEYWORDS: Akaike Information Criterion (AIC), Cox Proportional Hazard Model, Heart Failure, Parametric Models.

1. INTRODUCTION

The last stage of many heart diseases is heart failure, a complex condition. Significant mortality, poorer quality of life, and high hospitalization rates are the results. Making clinical decisions, evaluating risk, and creating focused treatments all depend on knowing and forecasting how long heart failure patients will live. Survival analysis requires statistical models. They offer information about the likelihood of significant events, such as hospital readmissions or deaths.

One of the most popular techniques in survival analysis is the Cox Proportional Hazards model, which was first presented by Sir David Cox in 1972. Without requiring the baseline hazard function to be specified, it provides estimates of hazard ratios associated with various factors. The Cox model is flexible and simple to understand because of its semi-parametric methodology. It is predicated on the idea that the risk ratios between groups don't change over time, though. Results could be skewed or deceptive if this presumption is broken.

On the other hand, parametric survival models assume specific distributions for survival times, such as the Exponential, Weibull, Log-logistic, Log-normal, Gamma, Gompertz, and Rayleigh

models are parametric models. When the assumed distribution fits the data well these parametric models can produce more accurate estimates and also allow direct estimation of hazard rates and survival functions. Parametric models can also manage developing hazards and forecast outcomes outside of the observed follow-up periods, both of which are frequently beneficial in clinical practice.

Given the advantages and disadvantages of each modeling technique, it is vital to understand how they work in various clinical situations. In this study, we use a dataset of heart failure patients to evaluate the Cox Proportional Hazards model with parametric survival models. We assess the models' fit, predictive power, and interpretability, and consider the implications for clinical research and practice.

2. LITERATURE REVIEW

Survival analysis is highly applied in clinical research to analyze time-to-event data, particularly for long-term diseases such as heart failure. Over the years, different modeling techniques have been developed and employed to further develop the knowledge of the factors influencing patient survival and facilitate improved clinical decision-making.

Cox Proportional Hazards (PH) model, introduced by Cox (1972), is used extensively due to its semi-parametric nature, without necessitating the specification of the baseline hazard function. The Cox model has been employed by several studies, including those by Levy et al. (2002), to identify clinical predictors of death in heart failure populations. While the model is powerful and interpretable, it will have a limitation in relying on the proportional hazards assumption, especially when covariate effects vary over time.

In contrast, parametric models assume a specific distribution for survival times and are especially useful when the data exhibit time-dependent hazard structures. The Weibull model, for example, allows for increasing or decreasing hazards over time and has been effectively used in several heart failure studies (e.g., D'Agostino et al., 2008). Similarly, Log-normal and Log-logistic models have been used to model more complex hazard shapes, especially when survival curves show non-monotonic behavior. Research by Klein and Moeschberger (2003) and Bradburn et al. (2003) emphasizes the advantages of parametric models in providing smooth and extrapolatable survival estimates.

Recent studies have explored model comparison frameworks to assess the relative performance of Cox and parametric models. A study by Royston and Parmar (2011) compared flexible parametric models with Cox models in clinical trials and found that parametric models often performed better in terms of predictive accuracy when their assumptions were met. Additionally, statistical tools such as the Akaike Information Criterion (AIC) and graphical methods have been used to guide model selection (Collett, 2015). Ravangard et al., (2011) compare the results of Cox proportional hazards model and parametric models in the study of length of stay in a Tertiary Teaching Hospital in Tehran, Iran. Pourhoseingholi et al., (2007) compare the Cox regression and parametric models for survival of patients with gastric carcinoma. Pourhoseingholi et al., (2009) used log-normal censored regression model to find out the prognostic factors in gastric cancer.

For heart failure, precise survival modeling is essential because of heterogeneity in the patient population. Research such as that presented by Pocock et al. (2006) has shown that considering time-dependent effects and flexible modeling strategies can very much improve prognostic

models. Yet few studies have directly compared parametric and semi-parametric survival models for heart failure cohorts systematically, and hence more research in this context is required. Ahmad et al. (2017) employed Cox regression model, Kaplan Meier plot and Martingale residuals to analysis of heart failure patients' survival. Chicco and Jurman (2020) utilized machine learning classifiers to analysis heart failure patients' survival. Ashine et al. (2021) utilized Cox proportional hazard model and Bayesian parametric survival models to analysis survival time of patients with heart failure.

3. DATASETS AND STATISTICAL TECHNIQUES

Dataset

We analyzed a dataset that included 299 heart failure patients' medical records that were gathered between April and December 2015 at the Allied Hospital and the Faisalabad Institute of Cardiology in Faisalabad, Punjab, Pakistan [Ahmad et al. (2017)]. The patients were in age from 40 to 95 years old, with 105 women and 194 males among them. Each of the 299 patients had a history of heart failure and left ventricular systolic dysfunction that classified them in heart failure stages III or IV according to the New York Heart Association's (NYHA) classification [Bredy et al. (2017)]. The dataset contains total 13 potential features which were described in Table 1. The features Age, CPK (Creatinine phosphokinase), Ejection Fraction (EF), Platelets, Serum creatinine, Serum sodium, Time (Follow-up period) are taken as continuous where the features Sex, Anemia, Blood pressure, Diabetes, Smoking, Event are taken as binary. The quantitative features of the dataset are presented in Table 2 and Table 3. Other additional details about this dataset can be found in [Ahmad et al. (2017)].

TABLE 1: DESCRIPTION OF EACH FEATURE OF THE DATASET

Categorical Variables		Continuous Variables	
Variables	Description (Numbers)	Variables	Description[Range]
Sex	0-Female (105); 1-Male (194)	Age	Age of the patient [40-95]
Anemia	0-Absence (170); 1-Presence (129)	CPK	Level of CPK enzyme in the blood[23-7861]
High Blood pressure (BP)	0-No (194); 1-Yes (105)	Ejection fraction (EF)	Percentage of blood leaving the heart at each contraction [14-80]
Diabetes	0-Absence (174); 1-Presence (125)	Platelets	Platelets in the blood in kiloplatelets/mL[25.01-850.00]
Smoking	0-No (203); 1-Yes (96)	Serum creatinine	Level of creatinine in the blood in mg/dL[0.50-9.40]
Event	0-Survived (203); 1-Deceased (96)	Serum sodium	Level of sodium in the blood in mEq/L[114-148]
		Time	Follow-up period in days [4-285]

TABLE 2: STATISTICAL QUANTITATIVE DESCRIPTION OF CATEGORY FEATURE OF THE DATASET

Variables	Full Sample (299)		Dead Patients (96)		Survived Patients (203)	
	Number	Percentage	Number	Percentage	Number	Percentage
Sex (0-Female)	105	35.12	34	35.42	71	34.98
Sex (1-Male)	194	64.88	62	64.58	132	65.02
Anemia (0-Absence)	170	56.86	50	52.08	120	59.11
Anemia (1-Presence)	129	43.14	46	47.92	83	40.89
High Blood pressure (0-No)	194	64.88	57	59.38	137	67.49
High Blood pressure (1-Yes)	105	35.12	39	40.62	66	32.51
Diabetes (0-Absence)	174	58.19	56	58.33	118	58.13
Diabetes (1-Presence)	125	41.81	40	41.67	85	41.87
Smoking (0-No)	203	67.89	66	68.75	137	67.49
Smoking (1-Yes)	96	32.11	30	31.25	66	32.51

TABLE 3: STATISTICAL QUANTITATIVE DESCRIPTION OF NUMERIC FEATURE OF THE DATASET

Variables	Full Sample (299)			Dead Patients (96)			Survived Patients (203)		
	Median	Mean	s.d.	Median	Mean	s.d.	Median	Mean	s.d.
Age	60.00	60.83	11.89	65.00	65.22	13.21	60.00	58.76	10.64
Ejection fraction	38.00	38.08	11.83	30.00	33.47	12.53	38.00	40.27	10.86
CPK	250.00	581.80	970.29	259.00	670.20	1316.58	245.00	540.10	753.80
Platelets	262.00	263.36	97.80	258.50	256.38	98.53	263.00	266.66	97.53
Serum sodium	137.00	136.60	4.41	135.50	135.40	5.00	137.00	137.20	3.98
Serum creatinine	1.10	1.39	1.03	1.30	1.84	1.47	1.00	1.19	0.65
Time	115.00	130.30	77.61	44.50	70.89	62.38	172.00	158.30	67.74

s.d.: standard deviation

Statistical Techniques

The Cox Proportional Hazards Model is a most widely used generalized and a powerful semi-parametric model used in survival analysis to investigate how different variables influence the probability of a certain event—like failure, relapse, or death—happening. The model displays the following hazard function

$$h(t|X) = h_0(t) \cdot \exp(\beta_1 X_1 + \beta_2 X_2 + \dots + \beta_p X_p)$$

where $h(t|X)$ is the hazard function for the covariate X at time t , $h_0(t)$ is the baseline hazard function which does not need to be specified, X_1, X_2, \dots, X_p are covariates and $\beta_1, \beta_2, \dots, \beta_p$ are the regression coefficients.

The model assumes that two single risk ratio is stable over a period of time:

$$\frac{h(t|X_1)}{h(t|X_2)} = \exp[\beta^T (X_1 - X_2)]$$

which is referred to as proportional hazard assumption. The regression coefficient β , positive indicates the hazard increases and thereby the probability of survive is reduced; negative indicates the hazard is reduced and thereby the probability of survive is improved. The hazard

ratio $(HR) \exp(\beta)$ shows by how much the risk increases or decreases if the covariate shifts by one unit.

An alternative to the Cox Proportional Hazards Model in survival analysis when proportional hazard assumption violated is the Aalen's Additive Hazard Model. Aalen's model implies that covariates have an additive effect on the hazard function, whereas Cox assumes that covariates have a multiplicative effect on the hazard. Also the Aalen's Additive Hazard Model is a non-parametric model which allows the effects of the covariates to change over time. The model is given by the following hazard function

$$h(t|X(t)) = \beta_0(t) + \beta_1(t)X_1(t) + \beta_2(t)X_2(t) + \dots + \beta_p(t)X_p(t)$$

where $h(t|X(t))$ is the hazard function for the covariate $X(t)$ at time t , $\beta_j(t)$ is the time-varying regression coefficient.

In addition, in case the proportional hazard assumption is bound to be in violation, the use of parametric survival models can prove better. Parametric models most commonly used are Exponential, Weibull, Log-logistic, Log-normal, Gamma, Gompertz and Rayleigh distributions.

Proportional hazards assumption was checked in this research by employing Schoenfeld residuals, Cox regression model and parametric models such as Exponential, Weibull, Log-logistic, Log-normal and Rayleigh. The p -value less than 0.05 are taken as statistical significant. Akaike Information Criterion (AIC) was used to compare the different model performance.

4. RESULTS AND DISCUSSION

The results for proportional hazards assumptions testing are shown in **Table 4** and a plot of Schoenfeld residuals for all the covariates is shown in **Figure 1**. From Table 4, the correlation between the Schoenfeld residuals for the variable 'Ejection.Fraction' and ranked survival time is -0.0277 with a p -value of 0.03. This significant p -values proof that the proportional hazards assumption is not satisfied for the variable 'Ejection.Fraction'. The p -values for the other variables are not significant suggest that there is not enough evidence to reject the proportional hazards assumptions for these variables. The global test for the entire model is not significant with $p = 0.39$. This global test offers evidence that the proportional hazards assumption is satisfied for that model.

TABLE 4: TEST FOR PROPORTIONAL HAZARD ASSUMPTIONS

Covariates	rho	chisq	p
Gender	-0.1054	0.0763	0.78
Smoking	0.0134	0.4790	0.49
Diabetes	0.0983	0.1920	0.66
BP	0.0074	0.0082	0.93
Anaemia	0.0840	0.0169	0.93
Age	0.2090	0.1030	0.75
Ejection.Fraction	-0.0277	4.6900	0.03*
Sodium	0.0728	0.1100	0.74
Creatinine	-0.0455	1.5200	0.22
Pletelets	0.1127	0.00006	1.00
CPK	-0.1140	1.02	0.31
GLOBAL		0.1170	0.39

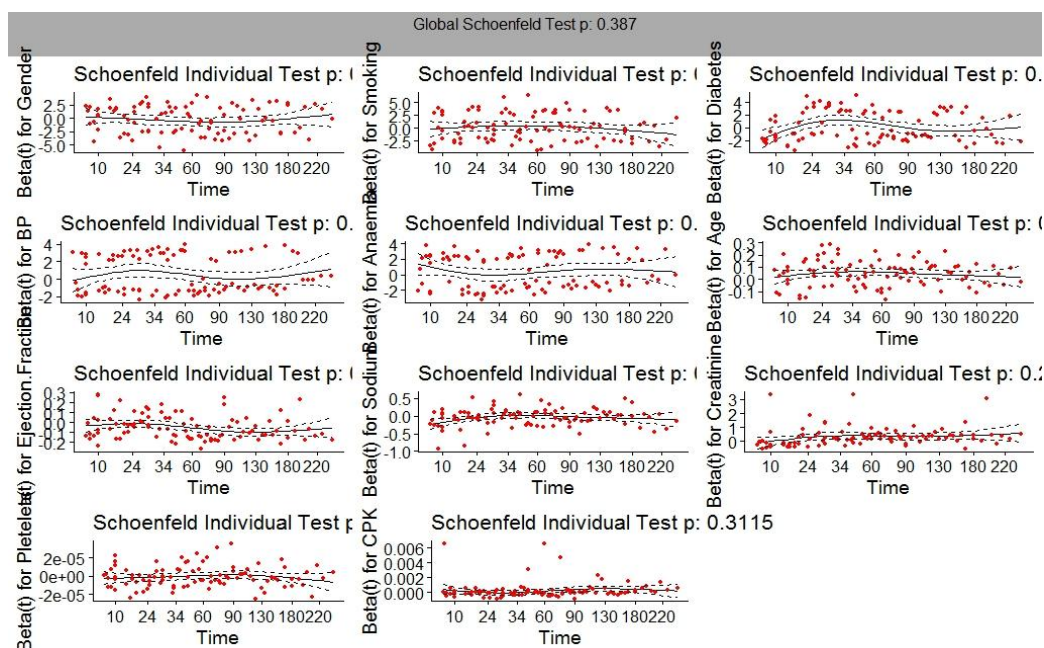


Figure 1. Schoenfeld residuals

The Cox proportional hazard model and the parametric models like, Exponential, Weibull, Log-logistic, Log-normal and Rayleigh was used separately to investigate the influence of several factors on the survival times.

The results of Cox regression model and the parametric models are presented in Table 5 – Table 10. The prognostic factors like Blood pressure (BP), Anaemia, Age, Ejection.Fraction, Creatinine and Creatinine phosphokinase (CPK) are the significant factors for survival of the heart failure patients as per all six models. Sodium is significant covariates as per log-normal and Rayleigh models. The performances of the models are compared through AIC values which are shown in Table 11. We see that Cox regression model has the lower AIC among all models and the parametric model exponential has the lowest AIC among all parametric models. Also comparisons of the results of the covariates between Cox and Exponential models are present in Table 12. Both models perform same to identify the effective risk factor for survival of the patient with heart failure.

TABLE 5: SIGNIFICANCE OF VARIABLES UNDER COX REGRESSION MODEL

Covariates	Coefficient	HR	Z-value	p-value
Gender	-0.2375	0.7886	-0.944	0.3452
Smoking	0.1289	1.1376	0.513	0.6078
Diabetes	0.1399	1.1501	0.627	0.5307
BP	0.4757	1.6092	2.201	0.0278*
Anaemia	0.4601	1.5843	2.122	0.0338*
Age	0.0464	1.0475	4.977	6.45e-07*
Ejection.Fraction	-0.0489	0.9522	-4.672	2.98e-06*
Sodium	-0.0442	0.9568	-1.899	0.0575
Creatinine	0.3210	1.3786	4.575	4.76e-06*
Platelets	-4.635e-07	1.0000	-0.412	0.6806
CPK	2.207e-04	1.0002	2.225	0.0260*

TABLE 6: SIGNIFICANCE OF VARIABLES UNDER EXPONENTIAL MODEL

Covariates	Coefficient	SE	Z-value	p-value
Gender	0.234	0.252	0.93	0.352
Smoking	-0.118	0.251	-0.47	0.639
Diabetes	-0.142	0.223	-0.64	0.524
BP	-0.507	0.214	-2.38	0.017*
Anaemia	-0.492	0.214	-2.29	0.022*
Age	-0.0486	0.0093	-5.23	1.7e-07*
Ejection.Fraction	0.0509	0.0106	4.82	1.4e-06*
Sodium	0.0437	0.0231	1.90	0.058
Creatinine	-0.325	0.0681	-4.77	1.8e-06*
Pletelets	5.16e-07	1.13e-06	0.45	0.649
CPK	-2.38e-04	9.95e-05	-2.39	0.017*

TABLE 7: SIGNIFICANCE OF VARIABLES UNDER WEIBULL MODEL

Covariates	Coefficient	SE	Z-value	p-value
Gender	0.246	0.263	0.93	0.350
Smoking	-0.119	0.261	-0.45	0.649
Diabetes	-0.147	0.232	-0.63	0.528
BP	-0.514	0.222	-2.31	0.021*
Anaemia	-0.5	0.224	-2.24	0.025*
Age	-0.0498	0.01	-4.96	7.2e-07*
Ejection.Fraction	0.0525	0.0116	4.53	5.9e-06*
Sodium	0.0450	0.0241	1.87	0.062
Creatinine	-0.333	0.0731	-4.56	5.1e-06*
Pletelets	5.51e-07	1.18e-06	0.47	0.641
CPK	-2.43e-04	1.04e-04	-2.34	0.019*

TABLE 8: SIGNIFICANCE OF VARIABLES UNDER LOG-LOGISTIC MODEL

Covariates	Coefficient	SE	Z-value	p-value
Gender	0.256	0.285	0.90	0.369
Smoking	-0.123	0.283	-0.44	0.663
Diabetes	-0.136	0.248	-0.55	0.585
BP	-0.522	0.247	-2.12	0.034*
Anaemia	-0.489	0.245	-1.99	0.046*
Age	-0.0508	0.0105	-4.85	1.2e-06*
Ejection.Fraction	0.0506	0.0122	4.14	3.5e-05*
Sodium	0.0516	0.0270	1.91	0.056
Creatinine	-0.360	0.0979	-3.67	0.0002*
Pletelets	6.77e-07	1.28e-06	0.53	0.597
CPK	-2.36e-04	1.16e-04	-2.03	0.042*

TABLE 9: SIGNIFICANCE OF VARIABLES UNDER LOG-NORMAL MODEL

Covariates	Coefficient	SE	Z-value	p-value
Gender	0.176	0.296	0.59	0.5520
Smoking	-0.0862	0.292	-0.30	0.7675
Diabetes	-0.0864	0.255	-0.34	0.7343
BP	-0.503	0.256	-1.97	0.0489*
Anaemia	-0.524	0.253	-2.07	0.0384*
Age	-0.0482	0.0107	-4.51	6.5e-06*
Ejection.Fraction	0.0443	0.0115	3.84	0.0001*
Sodium	0.0608	0.0269	2.26	0.0239*
Creatinine	-0.359	0.0104	-3.44	0.0006*
Pletelets	7.16e-07	1.32e-06	0.54	0.5880
CPK	-2.55e-04	1.17e-04	-2.18	0.0291*

TABLE 10: SIGNIFICANCE OF VARIABLES UNDER RAYLEIGH MODEL

Covariates	Coefficient	SE	Z-value	p-value
Gender	0.0727	0.128	0.57	0.5686
Smoking	-0.0837	0.127	-0.66	0.5107
Diabetes	-0.0767	0.114	-0.67	0.5002
BP	-0.401	0.107	-3.74	0.0002*
Anaemia	-0.376	0.108	-3.49	0.0005*
Age	-0.0321	0.0049	-6.47	9.5e-11*
Ejection.Fraction	0.0306	0.0054	5.61	2.1e-08*
Sodium	0.0272	0.0116	2.34	0.0191*
Creatinine	-0.213	0.0361	-5.90	3.7e-09*
Pletelets	7.85e-08	5.87e-07	0.13	0.8935
CPK	-1.72e-04	5.26e-05	-3.27	0.0011*

TABLE 11: MODEL COMPARISON AS PER AIC

Models	Cox	Exponential	Weibull	Log-logistic	Log-normal	Rayleigh
AIC	958.46	1280.42	1282.24	1285.46	1287.37	1374.36

TABLE 12: COMPARISON OF THE RESULTS OF THE COVARIATES BETWEEN COX AND EXPONENTIAL MODELS

Covariates	Cox model			Exponential model		
	Coefficient	SE	p-value	Coefficient	SE	p-value
Gender	-0.2375	0.2516	0.3452	0.234	0.252	0.352
Smoking	0.1289	0.2512	0.6078	-0.118	0.251	0.639
Diabetes	0.1399	0.2231	0.5307	-0.142	0.223	0.524
BP	0.4757	0.2162	0.0278*	-0.507	0.214	0.017*
Anaemia	0.4601	0.2168	0.0338*	-0.492	0.214	0.022*
Age	0.0464	0.0093	6.45e-07*	-0.0486	0.0093	1.7e-07*
Ejection.Fraction	-0.0489	0.0105	2.98e-06*	0.0509	0.0106	1.4e-06*
Sodium	-0.0442	0.0233	0.0575	0.0437	0.0231	0.058
Creatinine	0.3210	0.0702	4.76e-06*	-0.325	0.0681	1.8e-06*
Pletelets	-4.635e-07	1.13e-06	0.6806	5.16e-07	1.13e-06	0.649
CPK	2.207e-04	9.92e-05	0.0260*	-2.38e-04	9.95e-05	0.017*

5. CONCLUSION

This research offers a comparative assessment of the Cox Proportional Hazards model and widely applied parametric survival models in the analysis of patient survival with heart failure. The Cox model, due to its semi-parametric flexibility and interpretability, is still an influential model, particularly when the proportional hazards assumption can be made. Nevertheless, parametric models like the Weibull and Log-normal provide useful alternatives, especially when survival times are known to follow established distributions or when extrapolation beyond the observed region is needed.

Our results show that the prognostic variables such as Blood pressure (BP), Anaemia, Age, Ejection.Fraction, Creatinine and Creatinine phosphokinase (CPK) are the important factors for survival of the heart failure patients according to all six models. Sodium is important covariates according to log-normal and Rayleigh models. Parametric models are capable of outperforming the Cox model in model fit and predictive accuracy when their assumptions are met.

Finally, the selection between Cox and parametric models ought to be determined by the nature of the data, the clinical question being investigated, and diagnostic tests like residual analysis and goodness-of-fit tests. The implementation of both modeling methods may give a better insight into the survival of patients and the strength of clinical conclusions obtained from survival analysis.

REFERENCES

1. Cox, D.R. (1972). Regression Models and Life-Tables (with discussion). *Journal of the Royal Statistical Society: Series B*, 34(2), 187–220.
2. Levy D., Kenchaiah S., Larson M. G., Benjamin E. J., Kupka M. J., Ho K. K. L., Murabito J. M., & Vasan R. S. (2002). Long-term trends in the incidence of and survival with heart failure. *The New England Journal of Medicine*, 347(18), 1397-1402.
3. D'Agostino R.B., Vasan R.S., Pencina M.J, Wolf P.A., Cobain M., Massaro J.M., Kannel W.B. (2008). General Cardiovascular Risk Profile for Use in Primary Care. *Circulation*, 117(6), 743-753.
4. Klein, J.P., Moeschberger, M.L. (2003). *Survival Analysis: Techniques for Censored and Truncated Data*. Springer.
5. Bradburn, M.J., Clark, T.G., Love, S.B., Altman, D.G. (2003). *Survival Analysis Part II: Multivariate data analysis – an introduction to concepts and methods*. *British Journal of Cancer*, 89(3), 431–436.
6. Royston, P., Parmar, M.K. (2011). Flexible parametric proportional hazards and proportional odds models for censored survival data, with application to prognostic modelling and estimation of treatment effects. *Statistics in Medicine*, 31(21), 2176–2197.
7. Collett, D. (2015). *Modelling Survival Data in Medical Research*. CRC Press.
8. Ravangard, R., Arab, M., Rashidian, A., Akbarisari, A., Zare, A., Zeraati, H. (2011). Compare the results of Cox proportional hazards model and parametric models in the study of length of stay in a Tertiary Teaching Hospital in Tehran, Iran. *Acta Medica Iranica*, 49(10), 650-658.

9. Pourhoseingholi, M. A., Hajizadeh, E., Dehkordi, B. M., Saface, A., Abadi, A., Zali, M. R. (2007). Compare the Cox regression and parametric models for survival of patients with gastric carcinoma. *Asian Pacific Journal of Cancer Prevention*, 8, 412-416.
10. Pourhoseingholi, M. A., Dehkordi, B. M., Saface, A., Hajizadeh, E., Solhpour, A., Zali, M. R. (2009). Prognostic factors in gastric cancer using log-normal censored regression model. *Indian J Med Res* 129, March 2009, 262-267.
11. Ahmad, T., Munir, A., Bhatti, S. H., Aftab, M., Raza, M. A. (2017). Survival analysis of heart failure patients: A case study. *PLoS ONE* 12(7): e0181001. Doi: 10.1371/journal.pone.0181001. Dataset: https://plos.figshare.com/articles/Survival_analysis_of_heart_failure_patients_A_case_study/5227684/Accessed May 2025.
12. 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*, (2020) 20:16. Doi: 10.1186/s12911-020-1023-5.
13. Ashine, T., Muleta, G., Tadesse, K. (2021). Assessing survival time of heart failure patients: using Bayesian approach. *Journal of Big Data*, (2021) 8:156. Doi: 10.1186/s40537-021-00537-4.
14. Bredy C, Ministeri M, Kempny A, Alonso-Gonzalez R, Swan L, Uebing A, Diller G-P, Gatzoulis MA, Dimopoulos K. (2017). New York Heart Association (NYHA) classification in adults with congenital heart disease: relation to objective measures of exercise and outcome. *Eur Heart J – Qual Care Clin Outcomes*. 4(1):51–58.