BAREKENG: Journal of Mathematics and Its Applications
June 2026 Volume 20 Issue 2 Page 1527–1540
P-ISSN: 1978-7227 E-ISSN: 2615-3017
https://doi.org/10.30598/barekengvol20iss2pp1527-1540

SURVIVAL ANALYSIS OF CHRONIC KIDNEY FAILURE PATIENTS USING THE COX STRATIFIED MODEL AND RANDOM SURVIVAL FOREST

Assyifa Lala Pratiwi Hamid[1*], **Budi Susetyo**[2], **Anang Kurnia**[3]

[1,2,3]*School of Data Science, Mathematics, and Informatics, IPB University*
Jln. Raya Dramaga, Bogor, 16680, Indonesia

*Corresponding author's e-mail: * **assyifa.hamid@apps.ipb.ac.id***

Article Info

Article History:

Received: 18th June 2025
Revised: 13th July 2025
Accepted: 31st August 2025
Avaialable Online: 26th January 2026

Keywords:

Survival Analysis;
Cox Stratified Model;
Random Survival Forest;
Chronic Kidney Failure;
Mortality.

ABSTRACT

This study aims to analyze the factors influencing the survival of patients with chronic kidney failure undergoing hemodialysis and to compare the performance of the Cox Stratified Model and the Random Survival Forest (RSF) using retrospective data from 741 patients at Asy-Syifa General Hospital, Indonesia. To ensure model validity, a structured methodology was developed. First, continuous clinical variables were transformed into categorical or binary formats based on clinical thresholds. The Cox model was initially applied, and key assumptions—linearity, multicollinearity, and proportional hazards—were tested. Violations in the proportional hazard assumptions were addressed by stratifying the Cox model based on hemodialysis frequency and hypertension status. Additionally, age was categorized to correct non-linearity. In parallel, RSF was trained to capture non-linear patterns and complex interactions without relying on such assumptions. The performance of both models was evaluated using the Concordance Index (C-Index), yielding 0.66 for the Cox Stratified Model and 0.6558 for RSF, which are considered to represent moderate performance. Key predictors of survival identified in the analysis included patient age, presence of hypertension, diabetes status, anemia, and frequency of hemodialysis sessions. Study limitations encompass the retrospective design confined to a single center, as well as interpretability challenges inherent to the Random Survival Forest (RSF) methodology. The originality of this research lies in its methodological rigor and the direct comparison between a statistically extended Cox model and a machine learning approach, offering practical insights for improving clinical risk stratification in chronic kidney failure management in Indonesia.

How to cite this article:

A. L. P. Hamid, B. Susetyo and A. Kurnia., "SURVIVAL ANALYSIS OF CHRONIC KIDNEY FAILURE PATIENTS USING THE COX STRATIFIED MODEL AND RANDOM SURVIVAL FOREST," *BAREKENG: J. Math. & App.*, vol. 20, no. 2, pp. 1527-1540, Jun, 2026.

Journal homepage: *https://ojs3.unpatti.ac.id/index.php/barekeng/*
Journal e-mail: *barekeng.math@yahoo.com; barekeng.journal@mail.unpatti.ac.id*
Research Article · Open Access

1. INTRODUCTION

Survival analysis is a statistical approach for data analysis to examine the time until an event occurs, particularly in medical and reliability studies [1]. This method enables researchers to analyze time-to-event data, providing insights into the duration until events such as failure, relapse, or death occur. The approach is fundamental in fields where understanding the timing of events is crucial for prognosis and decision-making. The Cox Proportional Hazards (Cox PH) model remains the most widely applied method. This model relies on the assumption of proportional hazards, which means that the effect of covariates on the hazard rate is constant over time. However, this assumption is often violated in real-world clinical data, where the influence of risk factors may vary at different time points, challenging the model's applicability and accuracy [2].

To address this limitation, extensions of the Cox model and machine learning-based survival models have emerged [3]. The Cox Stratified Model extends the standard Cox regression by allowing the baseline hazard to vary across predefined strata, making it suitable for handling variables that violate the proportional hazards assumption and for accounting for subgroup differences in baseline risk [4]. By stratifying variables with non-proportional hazards, this model preserves the proportionality assumption for other covariates while improving model validity. Meanwhile, machine learning models, such as Random Survival Forest (RSF), offer a flexible alternative that does not rely on strict assumptions [5]. RSF employs ensemble learning through decision trees to capture complex, non-linear interactions and handle high-dimensional data, making it particularly useful for healthcare datasets with mixed variable types and censoring.

Previous studies have shown varying performances of these models across different domains. For instance, [6] found RSF to outperform Survival Support Vector Machine (SSVM) in retail customer survival prediction, while Spreafico et al. [7] reported that the Cox model and Machine Learning have similar performance, but the extended Cox model gave better performance. Each method offers distinct advantages: the Cox Stratified Model retains the interpretability of hazard ratios and is well-suited for clinical decision-making, especially when proportional hazards assumptions can be met or addressed through stratification. It allows the separation of baseline hazards across strata, improving model validity while preserving the ability to assess covariate effects.

On the other hand, RSF offers a flexible, non-parametric approach capable of modelling complex, non-linear relationships and high-order interactions without relying on assumptions such as proportional hazards. It is particularly advantageous in high-dimensional clinical datasets where traditional models may fall short. In clinical settings, combining traditional statistical models with machine learning approaches is increasingly recommended to improve prediction accuracy and provide deeper insights into risk factors. Christiadi et al. [8] demonstrated the utility of Random Survival Forests for dynamic survival prediction in end-stage kidney disease, highlighting the method's flexibility in handling complex risk structures. Similarly, Sim et al. [9] compared Cox regression and machine learning approaches in a Malaysian CKD cohort and found that both methods produced competitive predictive performance, underscoring the value of combining interpretability from Cox models with the predictive strength of RSF.

This study aims to compare the Cox Stratified Model and RSF using empirical survival data from patients with chronic kidney failure at RSUD Asy-Syifa, Indonesia. The novelty of this research lies in its dual methodological approach, combining advanced statistical modeling with machine learning to evaluate patient survival in a clinical context where both linear assumptions and complex interactions coexist. Unlike prior studies that focus solely on either traditional or machine learning models, this study addresses key methodological challenges, such as proportional hazards violations and non-linearity by developing a tailored stratification strategy in the Cox model and leveraging the flexibility of RSF. Furthermore, the study provides one of the first direct comparisons of these methods in an Indonesian hemodialysis population, offering localized insights for clinical risk stratification. By examining and validating both models, this research seeks to determine significant survival predictors and identify the best-performing model for supporting clinical decision-making in managing chronic kidney failure patients.

2. RESEARCH METHODS

Survival analysis is a statistical method used to study the time until an event occurs, accommodating censored data where the event has not yet happened for some subjects [1]. A common challenge in survival data is handling right-censoring and time-dependent covariates.

2.1 Fundamental of Survival Analysis

The survival function $S(t)$ gives the probability beyond time t [10] written as Eq. (1):

$$S(t) = P(T \geq t) = 1 - F(t), \tag{1}$$

where $F(t)$ is the cumulative distribution function.

The hazard function $h(t)$ represent the instantaneous risk of the event at time t [10] written as Eq. (2):

$$h(t) = \lim_{\Delta t \to 0} \frac{P(t \leq T \leq t + \Delta t \mid T \geq t)}{\Delta t} = \frac{f(t)}{S(t)}, \tag{2}$$

where $f(t)$ is the probability density function of survival time. The relationship between survival and hazard function is [10] written as Eq. (3):

$$S(t) = exp\left(-\int_0^t h(u)\, du\right). \tag{3}$$

2.2 Cox Proportional Hazards Model and Stratified Cox Model

The Cox Proportional Hazards (Cox PH) model is a semi-parametric model that estimates the effect of covariates on survival time without requiring the specification of the baseline hazard function. The Cox Proportional Hazards (Cox PH) model is defined as [10] written as Eq. (4):

$$h(t|X) = h_0(t)\, \exp(\beta^T \mathrm{X}), \tag{4}$$

where $h_0(t)$ is the baseline hazard function, X is the covariate vector, and β is the coefficient vector. A key assumption of the Cox model is proportional hazards — meaning that the hazard ratios between groups are constant over time. This assumption is often violated in clinical datasets, especially when certain risk factors have time-varying effects [11].

To address this limitation, the Cox Stratified Model offers an extension by allowing the baseline hazard function to differ across strata (subgroups), while assuming proportional hazards within each stratum [12]. By stratifying on variables that violate the proportional hazards assumption, the Cox Stratified Model avoids biased estimates and retains the ability to estimate covariate effects for other variables. The Cox Stratified Model modifies the baseline hazard to vary across K strata [4] written as Eq. (5):

$$h(t|X, Z) = h_{0k}(t)\, \exp(\beta^T X), \tag{5}$$

where $h_{0k}(t)$ is the baseline hazard function specific to stratum k (defined by stratification variable Z), and the model β while allowing each stratum to have its own baseline hazard. The key assumptions of the Cox Stratified Model are:

1. The baseline hazard can vary by stratum but is unspecified (non-parametric).
2. The proportional hazards assumption still holds within each stratum.
3. Covariates not used for stratification have multiplicative effects on the hazard function.

Stratification is particularly useful when it affects the baseline hazard but does not need an explicit hazard ratio estimate. This makes the Cox Stratified Model a powerful tool when dealing with time-varying risks and non-proportional hazards, which are common in healthcare data [13].

2.3 Random Survival Forest

The Random Survival Forest (RSF) is a machine learning method designed to analyze right-censored survival data [14]. RSF is an ensemble method that builds multiple survival trees using bootstrapped samples and aggregates their predictions to improve accuracy and robustness. Unlike the Cox model, RSF makes no assumptions about the proportionality of hazards or the functional form of covariate effects. RF uses the log-

rank test to maximize the survival difference. Each tree estimates the Cumulative Hazard Function (CHF) using the Nelson-Aalen estimator [15] written as Eq. (6):

$$\hat{H}(t) = \sum_{t_i \le t} \frac{d_i}{r_i}, \tag{6}$$

where d_i is the number of events at the time t_i and r_i is the number at risk just prior to t_i. The ensemble CHF is obtained by averaging over B trees written as Eq. (7):

$$\hat{H}_{ensemble}(t|X) = \frac{1}{B}\sum_{b=1}^{B} \hat{H}_b(t|X). \tag{7}$$

Key advantages of RSF include:

1. Ability to capture non-linear and high-order interactions among covariates.
2. Handles both continuous and categorical variables naturally.
3. Robust to multicollinearity and does not require variable selection.
4. Provides variable importance measures and predicted cumulative hazard functions.

The RSF algorithm works by splitting nodes in each tree based on survival differences, often using the log-rank statistic. For each individual, the model aggregates the cumulative hazard estimates from all trees. It performs well in complex datasets where traditional models struggle due to non-linearities or violations of assumptions. However, RSF also has limitations, such as:

1. Lower interpretability compared to Cox models, since it does not directly provide hazard ratios.
2. Requires tuning of hyperparameters (e.g., number of trees, number of variables considered at each split).
3. Computationally more intensive, especially with large datasets.

Recent studies have highlighted RSF's effectiveness in predicting survival in diverse fields, including oncology, cardiovascular diseases, and customer churn [16]. RSF is especially useful when the goal is prediction rather than inference, and when the proportional hazards assumption cannot be made.

2.4 Model Evaluation

Model performance in survival analysis is commonly assessed using the Concordance Index (C-Index), which measures the model's discriminative ability [17]. The C-Index represents the proportion of all usable patient pairs where the model correctly predicts which patient has a longer survival time. Given N comparable pairs of individuals, the C-Index is calculated as [17] written as Eq. (8):

$$C = \frac{Number\ of\ concordant\ pairs + (0.5 \times Number\ of\ tied\ pairs)}{Number\ of\ comparable\ pairs}. \tag{8}$$

The C-Index ranges from 0.5 (no better than random chance) to 1.0 (perfect prediction). A higher C-Index indicates better concordance between predicted and actual outcomes.

1. C-Index ≈ 0.5: Poor model (random guess);
2. C-Index 0.6-0.7: Moderate performance;
3. C-Index > 0.7: Good predictive discrimination.

Both Cox-based models and machine learning models like RSF are evaluated using the C-Index to ensure comparability. RSF has been shown in some studies to achieve slightly higher C-Index values in datasets with complex interactions, while Cox models maintain interpretability and robustness [18]. Therefore, C-Index serves as a critical criterion in this study to compare the predictive performance of the Cox Stratified Model and Random Survival Forest in chronic kidney failure patients.

3. RESULTS AND DISCUSSION

This study utilized empirical survival data from 741 patients with chronic kidney failure undergoing hemodialysis at RSUD Asy-Syifa, Sumbawa Barat, recorded between January 2015 and December 2024. The dataset was originally collected from medical records and is classified as empirical *data*, containing both continuous and categorical clinical variables.

Prior to analysis, the raw data underwent a preprocessing step where selected variables were transformed into binary formats to facilitate model fitting and interpretation. The outcome variable was survival time (T), measured in days from the start of observation until either death or censoring. The event indicator (δ) was coded as 1 if death was attributable to chronic kidney failure and 0 if the patient was censored (due to other causes or loss to follow-up). A summary of the transformed variables is presented in Table 1.

Table 1. Variables

No	Variable	Definition	Values
1.	Status (δ)	Event indicator (death due to CKD)	1 = Death due to kidney failure, 0 = Death due to another reason or censored
2.	Survival Time (T)	Time from baseline to event or censoring (days)	Numeric
3.	Gender	Patient sex	1 = Male, 0 = Female
4.	Age	Patient age (years)	Numeric
5.	Hypertension	Based on systolic (≥140) or diastolic (≥90) pressure	1 = Yes, 0 = No
6.	Diabetes Mellitus	Based on fasting glucose (≥126 mg/dL)	1 = Yes, 0 = No
7.	Hemodialysis Frequency	Number of dialysis sessions per week	1 = ≥3 times/week, 0 = ≤2 times/week
8.	Anemia	Based on hemoglobin level (≤8 g/dL)	1 = Yes, 0 = No

This transformation process ensured that the dataset was suitable for both traditional survival models (such as the Cox Stratified Model) and machine learning models (such as Random Survival Forest), which require clearly defined covariates for accurate risk estimation.

3.1 Statistics Descriptive

Based on the data we used for this study, Table 2 show the statistics descriptive of the data:

Table 2. Statistics Descriptive of Chronical Kidney Patient in RSUD Asy-Syifa

Variable	Mean	Median	Std. Dev.	Min	Max
Time (Y)	837.8	783	543.5	3	2678
Status	0.73	-	0.44	0	1
Age (X1)	56.3	58	14.6	1	88
Gender (X2)	0.43	-	0.49	0	1
Systolic (X3)	137.4	134	21.6	90	180
Dyastolic (X3)	88.3	90	11.4	60	110
Diabetes Mellitus (X4)	121.7	118	26.8	70	200
Hemodialysis Frequency (X5)	2.46	3	0.65	1	3
Anemia (X6)	9.55	9.6	3.97	1	17

Data source: Result from the analysis conducted using Python

Table 2 presents the descriptive statistics of chronic kidney disease patients treated at RSUD Asy-Syifa. The average survival time of patients was 837.8 days with a median of 783 days, ranging widely from 3 to 2678

days, indicating substantial variation in survival duration. Approximately 73% of patients experienced the event of interest, while 27% were censored. The mean age of patients was 56.3 years (median 58), with an age range from 1 to 88 years, suggesting that most patients were middle-aged to elderly. Gender distribution showed that about 43% of the sample were male. The mean systolic blood pressure was 137.4 mmHg and the mean diastolic blood pressure was 88.3 mmHg, both values lying close to the standard hypertension thresholds, indicating that a considerable proportion of patients were hypertensive.

The mean fasting blood glucose level was 121.7 mg/dL (median 118), with a maximum value reaching 200 mg/dL, suggesting that diabetes mellitus was also prevalent in this population. On average, patients underwent hemodialysis 2.46 times per week, with a median of 3 sessions, reflecting that many were on an intensive treatment regimen. The mean hemoglobin level was 9.55 g/dL, below the normal reference range, indicating a high prevalence of anemia among patients, including some severe cases. Overall, these descriptive statistics highlight the heterogeneity of the patient population, with most individuals presenting comorbid conditions such as hypertension, diabetes, and anemia, alongside the need for frequent hemodialysis.

3.2 Cox Proportional Hazards Model

The purpose of Cox Proportional Hazards Model is to estimate each variable's effect on survival time using Hazard Ratios (HR). Cox models have hypothesis, where:

Null Hypothesis H_0: The variable X has no significant effect on the hazard rate.

$$H_0 : \beta = 0$$

Alternative Hypothesis H_1: The variable X significantly affects the hazard rate.

$$H_1 : \beta \neq 0$$

Here's the result of Cox model shown in Table 3.

Table 3. Result of Cox Model in Chronic Kidney Disease Patient

Variabel	Coef	p-value	Interpretation Coef	HR (exp(coef))	Hazard Ratio Interpretation
Age (X1)	0.04	<0.005	Age significantly affects survival time.	1.04	Every additional year of age increases the risk of death by 4%.
Gender (X2)	-0.02	0.82	Gender doesn't significantly affect survival.	0.98	Males have a 2% lower risk than females, but not significant.
Hypertension (X3)	0.27	<0.005	Hypertension significantly affects survival time.	1.31	Patients with hypertension have a 31% higher risk of death.
Diabetes (X4)	0.27	<0.005	Diabetes significantly affects survival time.	1.32	Diabetes increases the risk of death by 32%.
Hemodialysis (X5)	0.44	<0.005	Hemodialysis frequency has significantly affects survival time.	1.55	Hemodialysis frequency ≥ 3 times increases the risk of death by 55%.
Anemia (X6)	0.24	0.01	Anemia significantly affects survival time.	1.27	Patients with severe anemia have a 27% higher risk of death.

Data source: Result from the analysis conducted using Python

After fitting the Cox model, we evaluated key assumptions, including linearity, the absence of multicollinearity, and the proportional hazards assumption. The PH assumption was tested using Schoenfeld residuals.

3.2.1 Linearity

Linearity was assessed using Martingale residuals. The result of the linearity test for the continuous variable is presented below. If the residuals show a random scatter around zero, the linearity assumption is considered to be satisfied [19]. Binary variables are assumed to meet the linearity assumption by default.

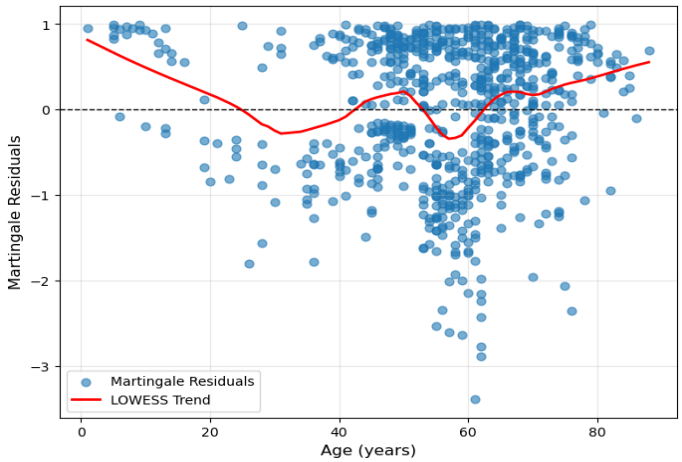

Figure 1. Result of Linearity Assumption Model in the Cox Model

The linearity assumption was assessed using Martingale residuals (Figure 1), which showed deviations suggesting non-linearity for age. Residuals tend to cluster around zero for ages roughly 40-80, but there is a curved pattern where residuals dip below zero at around age 60. There also appear to be some large negative residuals below -2, which might be outliers or suggest non-linearity. Since the Martingale residual plot indicated that the continuous variable Age (X1) did not satisfy the linearity assumption, we addressed this by transforming Age into a categorical variable [20]. Specifically, Age was divided into six categories: under 40 years, 40-49 years, 50-59 years, 60-69 years, 70-79 years, and 80 years or older. Due to a violation of the linearity assumption for the continuous age variable, Age was categorized into six intervals. The resulting categorical variable (Age_cat) was used in the Cox model, revealing a clear trend of increasing hazard with advancing age. Each category, compared to individuals younger than 40, demonstrated a statistically significant increase in the hazard ratio, with the peak observed in the 80+ group.

3.2.2 Multicollinearity

Variance Inflation Factor (VIF) is used to check for multicollinearity among your predictors, including categorical variables encoded as numbers. Table 4 shows the results of multicollinearity:

Table 4. Result of Multicollinearity Assumption Model in the Cox Model

Variables	VIF
Age (X1)	1.075
Gender (X2)	1.013
Hypertension (X3)	1.205
Diabetes (X4)	1.114
Hemodialysis (X5)	1.124
Anemia (X6)	1.080

Data source: Result from the analysis conducted using Python

Since the VIF value is below 5, there is no multicollinearity among the independent variables, and the assumption is met.

3.2.3 Schoenfeld Test for Proportional Hazards

The PH assumption is tested using Schoenfeld residuals. If all assumptions are satisfied—including the PH test—then the model is considered valid for interpretation. PH assumption has a hypothesis, where:

Null Hypothesis H_0: The hazard ratio for the covariate is proportional over time (i.e., it does not vary with time).

$$H_0: \frac{\partial}{\partial_t}\beta(t) = 0.$$

Alternative Hypothesis H_1: The hazard ratio for the covariate changes over time.

$$H_1: \frac{\partial}{\partial_t}\beta(t) = 0.$$

Table 5 shows the result of PH test:

Table 5. Result of Proportional Hazard Assumption Model in the Cox Model

Variable		P-Value	Interpretation
Age (X1)	40-49 yo	0.02	0.02 < 0.05 rejected H0, Violates PH assumption
	50-59 yo	0.06	0.06 > 0.05 accepting H0, PH assumption is hold
	60-69 yo	0.01	0.01 < 0.05 rejected H0, Violates PH assumption
	70-79 yo	0.01	0.01 < 0.05 rejected H0, Violates PH assumption
	>80 yo	0.02	0.02 < 0.05 rejected H0, Violates PH assumption
Gender (X2)		0.22	0.22 > 0.05 accepting H0, PH assumption is hold
Hypertension (X3)		0.93	0.93 > 0.05 accepting H0, PH assumption is hold
Diabetes (X4)		0.26	0.26 > 0.05 accepting H0, PH assumption is hold
Hemodialysis (X5)		0.01	0.01 < 0.05 rejected H0, Violates PH assumption
Anemia (X6)		0.1	0.1 > 0.05 accepting H0, PH assumption is hold

Data source: Result from the analysis conducted using Python

The age group 50–59 is the only one where the proportional hazards assumption holds. All other age categories violate the PH assumption**,** suggesting that their effect on survival changes over time. Hemodialysis (X5) violates the PH assumption — its impact on survival changes over time. All other covariates meet the PH assumption, and their effect on the hazard is stable over time. To resolve the PH violation in Hemodialysis, a stratified Cox model was employed using Hemodialysis (X5) as the stratifying variable. This allowed for separate baseline hazards among patients who underwent and did not undergo hemodialysis.

3.3 Cox Stratified Model

The Cox Stratified Model extends the standard Cox regression by allowing the baseline hazard to vary across predefined strata, making it particularly suitable for handling variables that violate the proportional hazards assumption. In this analysis, we stratified by hemodialysis frequency, since the PH assumption was violated for this variable. Stratification allows each hemodialysis group to have its own baseline hazard function, while still estimating the effects of other covariates across the entire sample. Table 6 shows the result of the Cox stratified model:

Table 6. Cox Stratified Model Result by Hemodialysis Frequency

Variabel		Coef	p-value	Interpretation Coef	HR (exp(coef))	Hazard Ratio Interpretation
Age (X1)	40-49 yo	1.14	<0.005	Age significantly affects survival time.	3.12	Patients aged 40–49 have 3.12× higher risk of death than those under 40.

Variabel	Coef	p-value	Interpretation Coef	HR (exp(coef))	Hazard Ratio Interpretation
50-59 yo	0.89	<0.005		2.43	Patients aged 50–59 have a 2.43× higher risk of death.
60-69 yo	1.65	<0.005		5.21	Patients aged 60–69 have a 5.21× higher risk of death.
70-79 yo	2.18	<0.005		8.81	Patients aged 70–79 are at 8.81× higher risk of death.
>80 yo	2.73	<0.005		15.3	Patients aged over 80 have the highest risk, with a 15.3× greater likelihood of death compared to those under 40.
Gender (X2)	-0.01	0.93	Gender doesn't significantly affect survival.	0.99	Males have a 1% lower risk than females, but not significant.
Hypertension (X3)	0.26	0.01	Hypertension significantly affects survival time.	1.29	Patients with hypertension have a 29% higher risk of death.
Diabetes (X4)	0.27	<0.005	Diabetes significantly affects survival time.	1.31	Diabetes increases the risk of death by 31%.
Anemia (X6)	0.24	0.01	Anemia significantly affects survival time.	1.27	Patients with severe anemia have a 27% higher risk of death.

Data source: Result from the analysis conducted using Python

After stratifying the model by hemodialysis frequency, age remained a significant predictor of survival. The hazard ratio increased with age, indicating that older patients faced a progressively higher risk of mortality. All age groups showed statistically significant results ($p < 0.005$), reinforcing age as a critical determinant of survival in patients with chronic kidney disease.

After stratifying the Cox model by hemodialysis frequency, we re-tested the Proportional Hazards (PH) assumption using the Schoenfeld residual test. This step is crucial to ensure that the remaining covariates in the stratified model satisfy the PH assumption. Stratification removes the need to model the baseline hazard for the stratified variable (hemodialysis), but it does not automatically guarantee that other variables meet the PH assumption. Table 7 shows the result of the PH test:

Table 7. PH Test Result of Cox Stratified Model

Variable		P-Value	Interpretation
Age (X1)	40-49 yo	0.99	0.99 > 0.05 accepting H0, PH assumption is hold
	50-59 yo	0.87	0.87 > 0.05 accepting H0, PH assumption is hold
	60-69 yo	0.88	0.88 > 0.05 accepting H0, PH assumption is hold
	70-79 yo	0.15	0.15 > 0.05 accepting H0, PH assumption is hold
	>80 yo	0.54	0.54 > 0.05 accepting H0, PH assumption is hold
Gender (X2)		0.81	0.81 > 0.05 accepting H0, PH assumption is hold
Hypertension (X3)		<0.005	<0.05 rejected H0, violates PH assumption

Variable	P-Value	Interpretation
Diabetes (X4)	0.94	0.94 > 0.05 accepting H0, PH assumption is hold
Anemia (X6)	0.46	0.46 > 0.05 accepting H0, PH assumption is hold

Data source: Result from the analysis conducted using Python

All variables except hypertension satisfy the Proportional Hazards (PH) assumption. Since Hypertension (X3) violates the PH assumption ($p < 0.05$), it may require further stratification or modeling adjustments (e.g., including time-dependent covariates or interaction terms). To solve these issues, we combine both hemodialysis frequency and hypertension variable for stratification. Table 8 shows the final result of Cox Stratified Model:

Table 8. Cox Stratified Model Result by Hemodialysis Frequency and Hypertension

Variabel		Coef	p-value	Interpretation Coef	HR (exp(coef))	Interpretasi Hazard Ratio (HR)
Age (X1)	40-49 yo	1.10	<0.005	Age significantly affects survival time.	3.01	Patients aged 40–49 have 3.01× higher risk of death than those under 40.
	50-59 yo	0.86	<0.005		2.36	Patients aged 50–59 have a 2.36× higher risk of death.
	60-69 yo	1.63	<0.005		5.10	Patients aged 60–69 have a 5.1× higher risk of death.
	70-79 yo	2.20	<0.005		9.05	Patients aged 70–79 are at 9.05× higher risk of death.
	>80 yo	2.55	<0.005		12.79	Patients aged over 80 have the highest risk, with a 12.79× greater likelihood of death compared to those under 40.
Gender (X2)		-0.04	0.64	Gender doesn't significantly affect survival.	0.96	Males have a 4% lower risk than females, but not significant.
Diabetes (X4)		0.28	<0.005	Diabetes significantly affects survival time.	1.32	Diabetes increases the risk of death by 32%.
Anemia (X6)		0.30	<0.005	Anemia significantly affects survival time.	1.35	Patients with severe anemia have a 35% higher risk of death.

Data source: Result from the analysis conducted using Python

After stratifying the Cox model by hemodialysis frequency and hypertension, we re-tested the Proportional Hazards (PH) assumption using the Schoenfeld residual test, Table 9 shows the result.

Table 9. PH Test Result of Cox Stratified Model

Variable		P-Value	Interpretasi
Age (X1)	40-49 yo	0.60	0.60 > 0.05 accepting H0, PH assumption is hold
	50-59 yo	0.24	0.24 > 0.05 accepting H0, PH assumption is hold
	60-69 yo	0.08	0.08 > 0.05 accepting H0, PH assumption is hold
	70-79 yo	0.01	0.01 < 0.05 rejecting H0, PH assumption is slightly violates
	>80 yo	0.24	0.24 > 0.05 accepting H0, PH assumption is hold
Gender (X2)		0.94	0.94 > 0.05 accepting H0, PH assumption is hold
Diabetes (X4)		0.18	0.18 > 0.05 accepting H0, PH assumption is hold
Anemia (X6)		0.32	0.32 > 0.05 accepting H0, PH assumption is hold

Data source: Result from the analysis conducted using Python

All remaining variables passed the PH test, except Age_cat[70–79], which showed a minor violation ($p = 0.01$). This minor violation was considered tolerable given the trade-off between interpretability and model complexity [21]. This suggests that the effect of being in the 70–79 age group on the hazard of death may change slightly over the follow-up period, rather than remaining constant. In practical terms, this could mean that the relative risk associated with this age group either increases or decreases over time due to interactions with unmeasured factors such as disease progression rates, comorbidity burden, or treatment adherence patterns.

Although this deviation from the PH assumption is statistically significant, its magnitude is likely small given that all other age categories met the assumption and the model's overall fit remained robust. In clinical survival analysis, such minor violations are often considered acceptable, especially when the trade-off between model complexity and interpretability is taken into account. Alternative approaches, such as modeling the age group as a time-dependent covariate, could address this issue but would add complexity without necessarily improving predictive accuracy in this dataset. Given the model's strong performance (C-Index 0.66) and the interpretability benefits of retaining categorical age variables, the minor violation was deemed tolerable for the purposes of this study.

3.4 Random Survival Forest

The Random Survival Forest (RSF) model was applied to explore complex, non-linear relationships and interactions between covariates without relying on the proportional hazard assumption [22]. The RSF was implemented using the RandomSurvivalForest class from the scikit-survival library with the following hyperparameters: 100 trees (n_estimators=100), random subset of predictors per split (max_features="sqrt"), minimum of 10 samples to split a node (min_samples_split=10), and minimum of 5 samples per leaf (min_samples_leaf=5). The number of jobs was set to 1 for consistent execution order, and a fixed random seed (random_state=42) ensured reproducibility. Hyperparameter values were chosen based on prior literature to balance bias and variance, and no grid search was performed because the study's focus was methodological comparison rather than hyperparameter optimization The model was trained using six predictor variables: age, gender, hypertension status, diabetes status, hemodialysis frequency, and anemia status. After splitting the dataset into training and testing sets (80:20), the RSF model was fitted using 100 trees.

The model's performance was evaluated using the Concordance Index (C-Index), a measure of discriminative ability. The resulting C-Index was 0.6558, indicating a moderate predictive performance. Although RSF does not provide direct hazard ratios, it is particularly useful for identifying complex patterns in survival data, especially when classical assumptions such as proportional hazards are not met. However, the limitation in interpretability can be a challenge in clinical practice, where clinicians often prefer models that quantify risk in terms of hazard ratios for clear decision-making. To address this, post-hoc explainability tools such as variable importance measures, partial dependence plots, or accumulated local effects can be applied to RSF results, offering insights into how predictors contribute to survival outcomes without compromising predictive flexibility.

This result supports the use of RSF as a complementary approach to traditional Cox models, especially in real-world clinical data where assumptions may be violated. While Cox models remain advantageous for

interpretability, integrating of explainable machine learning techniques into RSF provides a pathway to making its results more clinically actionable.

3.5 Comparing Model Evaluation

Table 10 shows the C-Index comparison from both model:

Table 10. C-Index Comparison

Model	C-Index
Random Survival Forest	0.6558
Cox Stratified Model	0.66

To evaluate the predictive performance of different survival models, two main approaches were compared: the Cox Stratified Model and Random Survival Forest (RSF). The Cox Stratified Model achieved the highest Concordance Index (C-index) at 0.66, demonstrating its effectiveness when the proportional hazards assumption is appropriately addressed through stratification. This finding aligns with the study by [7], which reported that extended versions of the Cox model can outperform machine learning methods when key assumptions are met, and interpretability is essential in clinical decision-making. The interpretability of the Cox model lies in its ability to express the effects of covariates in terms of hazard ratios, which are easily understood in clinical contexts. For instance, in this study, patients aged 50–59 years exhibited a hazard ratio of 2.43, indicating that they faced more than twice the risk of death compared to patients under 40 years. Such straightforward measures allow clinicians to readily translate statistical results into practical guidance for patient management, whereas machine learning approaches, despite their predictive power, often function as "black boxes" and provide limited insight into the underlying risk structure. Furthermore, the Cox model remains advantageous because it not only estimates survival probabilities but also quantifies the relative impact of covariates, thereby offering clinically actionable insights.

On the other hand, RSF achieved a comparable C-index of 0.6558, indicating competitive predictive accuracy. Its strength lies in handling non-linear relationships and complex variable interactions without requiring the assumptions inherent in the Cox model. This result is consistent with findings from [6], where RSF outperformed other models in settings with non-linear patterns and high-dimensional data. RSF's robustness to assumption violations makes it particularly valuable in real-world clinical datasets, where such violations are common. However, it lacks the interpretability provided by hazard ratios, which may limit its use in practice-oriented clinical environments.

The close performance between these two methods in this study suggests that the choice of model should depend on the analytical objective: if interpretability and inference are prioritized, especially in healthcare contexts, the Cox Stratified Model is preferable. Conversely, when predictive accuracy in complex data structures is the main concern, RSF offers a flexible and assumption-free alternative. Thus, this study contributes to the growing body of literature that emphasizes the complementary use of traditional statistical models and machine learning techniques in survival analysis.

4. CONCLUSION

This study evaluated survival among patients with chronic kidney failure using both the Cox Stratified Model and Random Survival Forest (RSF). We found that age, hypertension, diabetes, anemia, and frequency of hemodialysis were significant predictors of patient survival. We found that age, hypertension, diabetes, anemia, and hemodialysis frequency were significant predictors of patient survival. The Cox Proportional Hazards model initially provided interpretable results through hazard ratios but violated key assumptions such as linearity and proportional hazards for some variables. To address this, we categorized age and implemented stratification for hemodialysis and hypertension, which successfully resolved assumption violations. The final Cox Stratified Model demonstrated strong interpretability and valid estimates, with a C-Index of 0.66, indicating good predictive performance. In contrast, the RSF model, which does not rely on strict assumptions, captured complex non-linear relationships and produced a comparable C-Index of 0.6558. Although RSF lacks direct interpretability, its flexibility makes it valuable in clinical contexts with complicated variable interactions. However, this study's single-center retrospective design may limit the

generalizability of its findings. Patient demographics, treatment protocols, and healthcare resources at RSUD Asy-Syifa may differ from those in other settings, which could potentially influence survival outcomes. In addition, reliance on historical medical records may introduce missing data, documentation inconsistencies, and unmeasured confounding. Multicenter validation in more diverse populations is needed to confirm the predictive performance of both the Cox Stratified Model and RSF.

In conclusion, the Cox Stratified Model remains an excellent choice when model assumptions are met and interpretability is a priority. Meanwhile, RSF provides a powerful alternative for predictive modeling when data complexity or assumption violations pose a challenge. Using both models in tandem can enhance decision-making in chronic kidney disease management by balancing statistical rigor and predictive strength.

Author Contributions

Assyifa Lala Pratiwi Hamid: Conceptualization, Methodology, Software, Data Curation, Formal Analysis, Writing—Original Draft, Visualization; Budi Susetyo: Supervision, Validation, Writing—Review and Editing, Project Administration; Anang Kurnia: Methodology, Validation, Writing—Review and Editing, Resources. All authors have read and approved the final manuscript and contributed significantly to the research design, interpretation of results, and manuscript preparation.

Funding Statement

This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors.

Acknowledgment

The authors would like to thank RSUD Asy-Syifa, Sumbawa Barat, for providing access to patient medical records that supported this study. We also express our gratitude to the Department of Statistics, IPB University, for technical and academic guidance throughout the research process.

Declarations

The authors declare that they have no competing interests.

Declaration of Generative AI and AI-assisted technologies

Generative AI tools (e.g., ChatGPT) were used solely for language refinement (grammar, spelling, and clarity). The scientific content, analysis, interpretation, and conclusions were developed entirely by the authors. The authors reviewed and approved all final text.

REFERENCES

[1] D. G. Kleinbaum and M. Klein, "STATISTICS FOR BIOLOGY AND HEALTH SURVIVAL ANALYSIS A SELF-LEARNING TEXT THIRD EDITION." [Online]. Available: http://www.springer.com/series/2848

[2] A. Sheng and S. K. Ghosh, "EFFECTS OF PROPORTIONAL HAZARD ASSUMPTION ON VARIABLE SELECTION METHODS FOR CENSORED DATA," *Stat Biopharm Res*, vol. 12, no. 2, pp. 199–209, Apr. 2020, doi: https://doi.org/10.1080/19466315.2019.1694578.

[3] Q. Tran Anh and K. Taniguchi, "COUPLING DYNAMICAL AND STATISTICAL DOWNSCALING FOR HIGH-RESOLUTION RAINFALL FORECASTING: CASE STUDY OF THE RED RIVER DELTA, VIETNAM," *Prog Earth Planet Sci*, vol. 5, no. 1, Dec. 2018, doi: https://doi.org/10.1186/s40645-018-0185-6.

[4] M. Reza Pahlevi and T. Wuryandari, "MODEL REGRESI COX STRATIFIED PADA DATA KETAHANAN," *JURNAL GAUSSIAN*, vol. 5, no. 3, pp. 455–464, 2016, [Online]. Available: http://ejournal-s1.undip.ac.id/index.php/gaussian

[5] Hong Wang and Gang Li, "A SELECTIVE REVIEW ON RANDOM SURVIVAL FORESTS FOR HIGH DIMENSIONAL DATA," *Quant Biosci*, vol. 36, no. 2, pp. 85–96, Nov. 2017, doi: https://doi.org/10.22283/qbs.2017.36.2.85.

[6] N. G. A. P. P. Suantari, A. Fitrianto, and B. Sartono, "COMPARATIVE STUDY OF SURVIVAL SUPPORT VECTOR MACHINE AND RANDOM SURVIVAL FOREST IN SURVIVAL DATA," *BAREKENG: Jurnal Ilmu Matematika dan Terapan*, vol. 17, no. 3, pp. 1495–1502, Sep. 2023, doi: https://doi.org/10.30598/barekengvol17iss3pp1495-1502.

[7] M. Spreafico, A. D. Hazewinkel, M. A. J. van de Sande, H. Gelderblom, and M. Fiocco, "MACHINE LEARNING VERSUS COX MODELS FOR PREDICTING OVERALL SURVIVAL IN PATIENTS WITH OSTEOSARCOMA: A RETROSPECTIVE ANALYSIS OF THE EURAMOS-1 CLINICAL TRIAL DATA," *Cancers (Basel)*, vol. 16, no. 16, Aug. 2024, doi: https://doi.org/10.3390/cancers16162880.

[8] D. Christiadi *et al.*, "DYNAMIC SURVIVAL PREDICTION OF END-STAGE KIDNEY DISEASE USING RANDOM SURVIVAL FORESTS FOR COMPETING RISK ANALYSIS," *Front Med (Lausanne)*, vol. 11, 2024, doi: https://doi.org/10.3389/fmed.2024.1428073.

[9] R. Sim, C. W. Chong, N. K. Loganadan, N. L. Adam, Z. Hussein, and S. W. H. Lee, "COMPARISON OF A CHRONIC KIDNEY DISEASE PREDICTIVE MODEL FOR TYPE 2 DIABETES MELLITUS IN MALAYSIA USING COX REGRESSION VERSUS MACHINE LEARNING APPROACH," *Clin Kidney J*, vol. 16, no. 3, pp. 549–559, Mar. 2023, doi: https://doi.org/10.1093/ckj/sfac252.

[10] F. Emmert-Streib and M. Dehmer, "INTRODUCTION TO SURVIVAL ANALYSIS IN PRACTICE," Sep. 01, 2019, *MDPI*. doi: https://doi.org/10.3390/make1030058.

[11] H. Zhao, W. Yao, X. Min, K. Gu, G. Yu, Z. Zhang, J. Cui, L. Miao, L. Zhang, X. Yuan, Y. Fang, X. Fu, C. Hu, X. Zhu, Y. Fan, Q. Yu, G. Wu, O. Jiang, X. Du, J. Liu and L. Zhang., "APATINIB PLUS GEFITINIB AS FIRST-LINE TREATMENT IN ADVANCED EGFR-MUTANT NSCLC: THE PHASE III ACTIVE STUDY (CTONG1706)," in *Journal of Thoracic Oncology*, Elsevier Inc., pp. 1533–1546, Sep. 2021, doi: https://doi.org/10.1016/j.jtho.2021.05.006.

[12] M. Therneau and P. M. Grambsch, "MODELING SURVIVAL DATA: EXTENDING THE COX MODEL," illustrated ed. New York, NY, USA: Springer, 2010

[13] D. J. Ratnaningsih, A. Saefuddin, A. Kurnia, and I. W. Mangku, "STRATIFIED-EXTENDED COX MODEL IN SURVIVAL MODELING OF NON-PROPORTIONAL HAZARD," in *IOP Conference Series: Earth and Environmental Science*, Institute of Physics Publishing, Jul. 2019, doi: https://doi.org/10.1088/1755-1315/299/1/012023.

[14] M. S. Zaenal, A. Fitrianto, and H. Wijayanto, "COMPARISON OF EXTREMELY RANDOMIZED SURVIVAL TREES AND RANDOM SURVIVAL FORESTS: A SIMULATION STUDY," *Scientific Journal of Informatics*, vol. 11, no. 3, pp. 635–644, Aug. 2024, doi: https://doi.org/10.15294/sji.v11i3.8464.

[15] K. L. Pickett, K. Suresh, K. R. Campbell, S. Davis, and E. Juarez-Colunga, "RANDOM SURVIVAL FORESTS FOR DYNAMIC PREDICTIONS OF A TIME-TO-EVENT OUTCOME USING A LONGITUDINAL BIOMARKER," *BMC Med Res Methodol*, vol. 21, no. 1, Dec. 2021, doi: https://doi.org/10.1186/s12874-021-01375-x.

[16] P. Saha., "PREDICTING TIME TO DIABETES DIAGNOSIS USING RANDOM SURVIVAL FORESTS," Feb. 07, 2024. doi: https://doi.org/10.1101/2024.02.03.24302304.

[17] E. Longato, M. Vettoretti, and B. Di Camillo, "A PRACTICAL PERSPECTIVE ON THE CONCORDANCE INDEX FOR THE EVALUATION AND SELECTION OF PROGNOSTIC TIME-TO-EVENT MODELS," *Academic Press Inc.,* Aug. 01, 2020, doi: https://doi.org/10.1016/j.jbi.2020.103496.

[18] A. Alabdallah, M. Ohlsson, S. Pashami, and T. Rögnvaldsson, "THE CONCORDANCE INDEX DECOMPOSITION: A MEASURE FOR A DEEPER UNDERSTANDING OF SURVIVAL PREDICTION MODELS," Feb. 2022, doi: https://doi.org/10.1016/j.artmed.2024.102781.

[19] R. W. Nahhas, M. Dehmer, and F. Emmert-Streib, "INTRODUCTION TO REGRESSION METHODS FOR PUBLIC HEALTH USING R," *Bookdown*, 2021. [Online]. Available: https://bookdown.org/rwnahhas/RMPH/survival-linearity.html

[20] Easy Med Stat, "LINEARITY ASSESSMENT IN COX REGRESSION ANALYSIS", *EasyMedStat*, Jul. 8, 2021.

[21] P. C. Austin, "STATISTICAL POWER TO DETECT VIOLATION OF THE PROPORTIONAL HAZARDS ASSUMPTION WHEN USING THE COX REGRESSION MODEL," *J Stat Comput Simul*, vol. 88, no. 3, pp. 533–552, Feb. 2018, doi: https://doi.org/10.1080/00949655.2017.1397151.

[22] A. B. M. Rezaei., "REVIEW OF RANDOM SURVIVAL FOREST METHOD," *J. Biostat. Epidemiol.*, vol. 6, no. 1, pp. 18–30, Nov. 2020.