INDONESIAN COUNCIL OF PREMIER STATISTICAL SCIENCE

Indonesian Council of Premier Statistical Science
ISSN: 3030-9956/https://dx.doi.org/10.24014/icopss.v4i2.37876

Volume 4 Issue 2, 51 – 57, August 2025

Survival Analysis of Kidney Failure Patients Using Nelson-Aalen Cumulative Hazard Function Estimation and Log-Rank Test

Nofita Dwi Safitri[1], Salwa Zahrotul Jannah[2]

Rado Institute Statistics Online – Riau, Indonesia
[2]Mathematics Study Program, Universitas Islam Negeri Sultan Syarif Kasim Raiu, Indonesia

Received: 23 July 2025 Revised: 27 August 2025 Accepted: 30 August 2025 Published: 31 August 2025

Abstract - *Kidney failure is a critical condition that requires intensive monitoring of the patient's survival. This study aims to analyze the survival of kidney failure patients using a survival analysis approach with Nelson-Aalen cumulative hazard function estimation and the Log-Rank Test. The data used consisted of 106 patients with kidney failure who underwent hospitalization for 35 days, with characteristic variables including disease severity (chronic or acute), sex, and age group (≤ 50 years or > 50 years). The results of Nelson-Aalen's estimation showed that patients with acute conditions, male, and > age 50 years had a visually higher risk of death. However, the results of the Log-Rank test showed that there was no significant difference in survival function between the three categories (p-value > 0.05). These findings suggest that although there are visual indications of a difference in risk, statistically, these clinical characteristics have not had a significant effect on patient survival in the observation period. This study suggests the need for further studies with data and a broader duration of observations to gain a deeper understanding.*

Keywords: *survival analysis, kidney failure, Nelson-Aalen, Log-Rank Test.*

1. Introduction

Kidney failure is a condition in which the kidneys lose their ability to optimally filter metabolic waste, thus having a serious impact on the patient's health and quality of life [1]. People with kidney failure, particularly chronic kidney failure (CKD), generally require long-term therapy such as hemodialysis or kidney transplantation to maintain survival. GGK is one of the serious global health problems, progressive and irreversible [2]. Based on the report of the World Health Organization (WHO), GGK accounts for around 850,000 deaths annually and ranks 12th as a cause of death and 17th as a cause of disability globally [3].

In Indonesia, the number of cases of kidney failure shows an increasing trend from year to year. Data from various hospitals show that the number of patients undergoing hemodialysis as a kidney replacement therapy continues to increase every year, reflecting the high burden of this disease [4]. The main risk factors for kidney failure include hypertension and diabetes mellitus, which significantly increase the likelihood of permanent kidney damage. In addition, various other factors such as old age, gender, smoking habits, the use of analgesic drugs and NSAIDs, and the consumption of energy supplement drinks also accelerate the deterioration of kidney function [5].

Seeing the increasing number of cases and the complexity of the risk factors, a deeper understanding of patient survival is becoming increasingly important. In this case, an analytical approach is needed that not only describes the event but is also able to evaluate the patient's risk and survival. One relevant statistical approach is survival analysis, which allows the measurement of a patient's survival time as well as the identification of factors that affect it. Thus, this information can be used to improve the effectiveness of clinical management as well as more targeted health program planning.

Analysis Survival: It is a statistical approach that is commonly used to evaluate the time until an event occurs, such as death or medical failure [6]. This method has the advantage of accommodating censored data (censored data), namely patient data that has not experienced an event until the end of the observation period, so that survival estimates become more accurate [7]. One of the methods used in the analysis of Survival is a Nelson-Aalen estimator, which serves to estimate the Hazard non-parametric cumulative and describes the accumulation of event risk over time [8]. In addition, statistical tests such as the Log-Rank Test are used to compare survival functions between groups of patients with different characteristics.

Previous research has applied a lot of survival analysis to study survival in a medical context. Research [4] analyzed data on patients with chronic kidney failure (CKD) undergoing hemodialysis at Dr. Soetomo Surabaya Hospital, and found that age, history of hypertension, and diabetes mellitus significantly affected the difference in survival between patient categories. Research [9] Using the Kaplan-Meier and Log-Rank test to analyze the effects of treatment D-penicillamine on patient survival in Primary Biliary Cirrhosis (PBC), and found that differences between treatment groups could be identified significantly. Furthermore, the research [10] compares the Estimators Kaplan-Meier and Nelson-Aalen on credit data, and shows that although the survival curve is similar, the Nelson-Aalen estimator generates value Hazard slightly higher cumulative. Then, the research [8] applies the Nelson-Aalen and Weighted Log-Rank Test to assess the survival of Type II Diabetes Mellitus patients, as well as to find that gender and dietary patterns have a significant effect on the patients' chances of survival. In addition, the research [11] applies the Kaplan-Meier and tests the Log-Rank to analyze the survival of school-age residents who drop out of school based on socioeconomic characteristics in rural and urban areas of Central Java. This study shows that the Kaplan-Meier effective for describing differences in survival between groups with different characteristic variables such as gender, parental education, and social assistance.

Based on research [4] and studies [8] that apply the Nelson-Aalen in estimating the cumulative hazard function, and are supported by information on the risk factors of [5]. The authors were interested in conducting a similar analysis on the data of kidney failure patients to see patterns in the risk of death over time. In addition, referring to studies [9] and [11] that utilize the Log-Rank Test in comparing survival functions between groups, the author also wants to test the difference in risk between categories on each factor, namely the severity of the disease, gender, and age group. Thus, this study is expected to provide a more comprehensive picture of the variation in risk patterns based on the clinical characteristics of patients.

2. Research Methods

2.1. *Data Description*

This study uses data *Survival* of 106 patients with kidney failure who underwent hospitalization during the 35-day observation period. The data used are secondary data obtained from scientific publications [12]. Each observation recorded the patient's treatment time (in days), the patient's status (censored or dead), and patient characteristics, which included disease severity (chronic or acute), gender (male or female), and age group (≤ 50 years or > 50 years). The details of the data can be seen in Table 1 below:

Table 1. Survival Data of Kidney Failure Patients Based on Time, Status, and Patient Characteristics

i	Treatment Time (ti)	n.risk (ni)	Status		Severity		Gender		Age	
					Chronic	Acutely	L	P	≤ 50	> 50
1	0	106	Sensor	2	1	1	1	1	2	0
			Die	0	0	0	0	0	0	0
2	1	104	Sensor	9	6	3	3	6	3	6
			Die	1	0	1	1	0	0	1
3	2	94	Sensor	13	6	7	6	7	6	7
			Die	3	1	2	2	1	1	2
4	3	78	Sensor	10	5	5	4	6	3	7
			Die	1	0	1	1	0	0	1
5	4	67	Sensor	6	4	2	5	1	4	2
			Die	2	2	0	0	2	1	1
6	5	59	Sensor	11	2	9	7	4	4	7
			Die	1	0	1	0	1	0	1
7	6	47	Sensor	8	0	8	4	4	2	6
			Die	0	0	0	0	0	0	0
8	7	39	Sensor	5	3	2	3	2	2	3
			Die	1	0	1	0	1	0	1
9	8	33	Sensor	4	1	3	2	2	3	1
			Die	0	0	0	0	0	0	0
10	9	29	Sensor	3	1	2	2	1	0	3
			Die	1	0	1	1	0	1	0
11	10	25	Sensor	1	0	1	0	1	0	1
			Die	1	0	1	1	0	0	1
12	11	23	Sensor	4	0	4	2	2	0	4
			Die	2	0	2	1	1	1	1
13	12	17	Sensor	1	0	1	1	0	0	1
			Die	1	0	1	1	0	0	1

14	13	15	Sensor	5	1	4	3	2	2	3
			Die	0	0	0	0	0	0	0
15	14	10	Sensor	0	0	0	0	0	0	0
			Die	1	0	1	1	0	0	1
16	17	8	Sensor	2	0	2	2	0	2	0
			Die	0	0	0	0	0	0	0
17	18	7	Sensor	1	0	1	1	0	1	0
			Die	1	0	1	0	1	0	1
18	20	5	Sensor	2	1	1	0	2	0	2
			Die	0	0	0	0	0	0	0
19	22	3	Sensor	1	0	1	0	1	0	1
			Die	0	0	0	0	0	0	0
20	28	2	Sensor	1	0	1	0	1	0	1
			Die	0	0	0	0	0	0	0
21	35	1	Sensor	1	0	1	0	1	0	1
			Die	0	0	0	0	0	0	0

Information:

ti : The time when an event or sensor occurs (in units of days).i

ni : Number at risk right ahead of time.ti

Status: Indicates whether the patient has a death event (*Event*) or is still alive at the time of observation end (*sensor*).

Severity: The patient's disease condition, categorized as chronic or acute.

Gender: Categorized as L (male) and P (female).

Age: Classified into ≤ 50 years and > 50 years.

2.2. Research Methods

This study uses *a survival* analysis approach to study the time to death in kidney failure patients. The methods used include estimation *of the cumulative* hazard function using *the Nelson-Aalen* method, estimation of the *survival* function, and the *Log-Rank Test* to compare survival curves between groups of patients based on certain characteristics. The stages of data analysis in this study are carried out as follows:

1. Describe the cumulative *hazard* curve using *the Nelson-Aalen method* based on the category of each variable.
2. Describe the *survival curve of* kidney failure patients based on characteristic categories.
3. Testing the differences in survival curves between groups using *the Log-Rank Test*.
4. Interpret the results descriptively according to the research objectives.

2.2.1. Survival Function

Random variable that declares the time of *Survival.* An object is usually denoted by and functions T*Survival* is noted with. This function indicates the probability that the object will still survive over time. Thus, the $S(t) t t > 0$*Survival* $S(t)$ can be defined as [13]:

$$S(t) = P(T > t) = 1 - P(T \leq t) = 1 - F(t) \quad (1)$$

If it is a random variable of an individual's lifetime in intervals, then its probability density function $T[0, \infty)$,*(probability density function*) is denoted with and its cumulative distribution function is $f(t) F(t)$[14]. *The probability density function* expresses the chances of failure (death) of an individual in a time interval and is formulated as:$(t, t + \Delta t)$

$$f(t) = \lim_{\Delta t \to 0} \frac{P(t \leq T < t + \Delta t)}{\Delta t} \quad (2)$$

The cumulative distribution function, which is the probability that an individual fails before or at the time, is defined as $F(t) t$[14]:

$$F(t) = P(T \leq t) = \int_0^t f(x) dx \quad (3)$$

where:

$F(t)$: Cumulative distribution function at time *t*,

$f(x)$: *Probability density function* at the time,x

$S(t)$: Function *Survival*.

2.2.2. Hazard Function

Function *Hazard:* It is defined as the rate of failure at the time, provided that the individual still survives until that time. This function expresses the probability of occurrence in time intervals per unit of time, assuming the individual lasts until time. Mathematically, the $t(t, t + \Delta t)t Hazard$ stated as [15]:

$$h(t) = \lim_{\Delta t \to 0} \frac{P(t \leq T < t + \Delta t \mid T \geq t)}{\Delta t} \quad (4)$$

The relationship between the *Survival* and the functions *Hazard* is [13]:

$$S(t) = \exp[-H(t)] \quad (5)$$

2.2.3. Nelson-Aalen Estimation

This method is also sometimes referred to as the empirical cumulative *Hazard* function, but more commonly known as the *Nelson-Aalen* (NA) estimator. The Nelson-Aalen method ***is*** used to estimate the *Hazard* cumulatively non-parametrically. Estimator *Nelson-Aalen* to *Hazard* cumulative at the time it is formulated as t[8]:

$$\widehat{H}(t) = \sum_{i \leq t} \frac{d_i}{n_i} \quad (6)$$

where:

$\widehat{H}(t)$: Estimation *Hazard* cumulative at the time,t

d_i : Number of events at the time,t_i

n_i : The number of individuals who are still at risk at the time t_i[16].

The survival function can be estimated from the following:$\widehat{H}(t)$

$$\hat{S}(w) = e^{-\widehat{H}(w)} \quad (7)$$

2.2.4. Log-Rank Test

The Log-Rank Test is used to compare the *survival* curves of two or more groups, with the aim of finding out if there is a significant difference between the curves. The hypotheses used in *the Log-Rank Test* are as follows:

H_0: There is no difference in *survival function* between groups.

H_1: There is at least one group that has different *survival* functions.

The test statistics used in *the Log-Rank Test* are divided into two, namely for two groups and for more than two groups. The test statistics for *the two-group Log-Rank Test* are as follows:

$$\chi^2 = \frac{(O_1 - E_1)^2}{Var\,(O_1 - E_1)} \quad (8)$$

For more than two groups, the test statistic *Log-Rank* can be calculated using [17]:

$$\chi^2 = \sum_{i=1}^{G} \frac{(O_i - E_i)^2}{E_i} \quad (9)$$

with:

G : Number of groups compared,

O_i : The number of events observed in the group,i

E_i : The expected number of events in the first group.i.

The Log-Rank Test generates a test statistical value that follows the chi-squared distribution**.** $\chi^2$with a degree of freedom. The statistical value of the test is used to calculate $G - 1$the p-value. Decision-making was carried out by comparing *the p-value* with the significance level of α = 0.05. If *the p-value* ≤ 0.05, it is rejected, which means that there is a significant difference between H_0 *the survival* functions. Conversely, if the *p-value* > 0.05, then it fails to be rejected.H_0

3. Analysis

3.1 Cumulative Hazard Curve

Figure 1 presents the cumulative hazard function curve based on *the Nelson-Aalen* estimate, which shows differences in mortality risk patterns according to three factors, namely disease severity, sex, and age group.

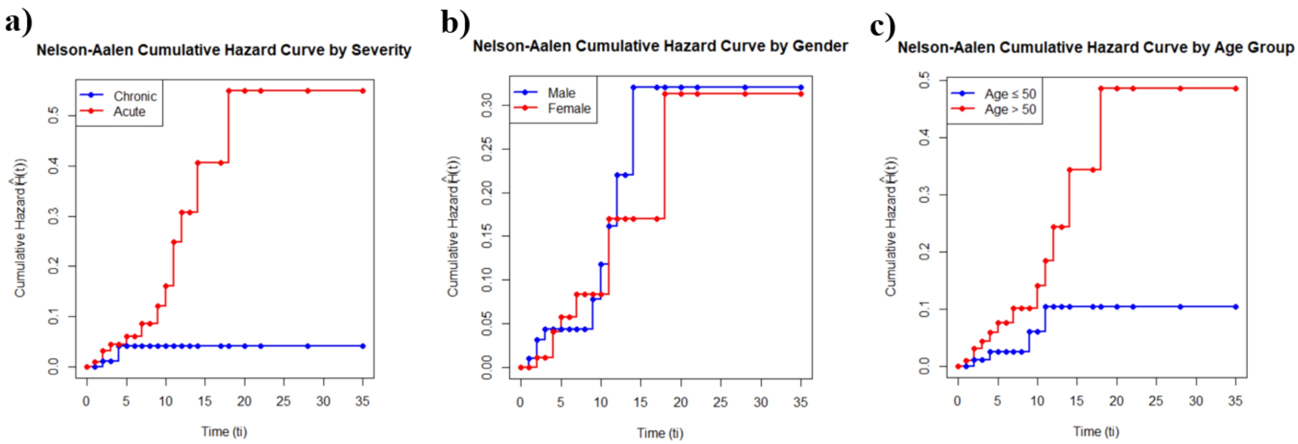

Figure 1. Cumulative Hazard Function Curve Using *the Nelson-Aalen* Method Based on Factors: (a) Disease Severity, (b) Gender, and (c) Age Group.

Based on Figure 1a, the cumulative *hazard* function curve in patients with acute conditions showed a faster and steeper increase in the risk of death compared to patients with chronic conditions, which showed a slower upward pattern and tended to be stable in later periods. This indicates that patients with acute conditions have a higher chance of dying in a shorter period of time, while patients with chronic conditions face a relatively lower risk. This difference in pattern confirms that the severity of the patient's condition greatly affects the risk of death during the observation period.

Furthermore, Figure 1b shows the difference in mortality risk patterns between male and female patients. The curve of male patients showed a sharper increase in the risk of mortality, especially after the 9th time, compared to female patients, who had a slower and more stable pattern of increase. This means that gender also influences the level of risk of death during the observation period, with male patients having a higher risk.

Then, Figure 1c shows the cumulative hazard function curve based on the age factor. The age group > 50 years experienced a faster and more consistent increase in the risk of death, with a steeper curve throughout the observation period. In contrast, the ≤ age group of 50 years showed a relatively flat and stable curve, especially after the 11th time, which signaled a lower risk of death in the younger age group. Thus, age is also an important factor in influencing the level of risk of death in kidney failure patients during the observation period.

3.2 *Survival Function Estimation Curve*

Figure 2 presents an estimate of the *survival* function based on *the Nelson-Aalen* approach for the same three factors. This curve represents the patient's survival probability during the observation period.

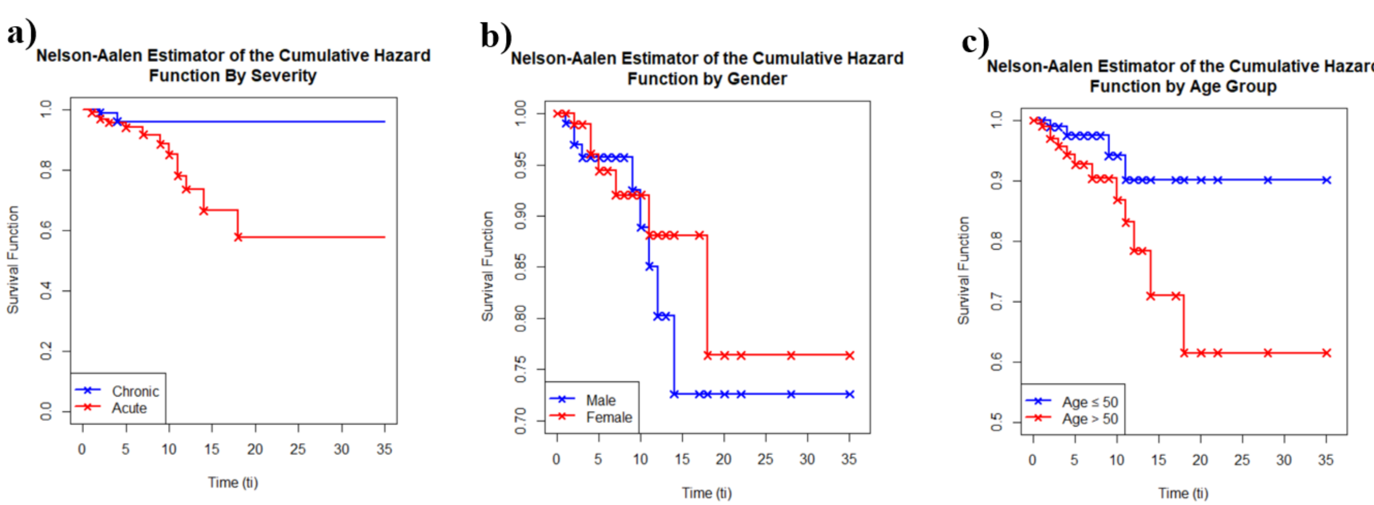

Figure 2. *Nelson-Aalen Survival* Function Estimation Curve Based on Factors: (a) Disease Severity, (b) Sex, and (c) Age Group

Figure 2a shows that patients with chronic conditions have a higher probability of survival compared to acute patients. The curve for acute patients appears to decline more sharply since the beginning of observation, indicating a higher risk of death in a shorter period of time. In contrast, patients with chronic conditions showed a more stable *survival* function curve and were close to 1 throughout the observation period. This indicates that the severity of the disease has a great influence on the patient's chances of survival.

Figure 2b shows the difference in the *survival* function by sex. The curve of male patients decreased more sharply, indicating a higher risk of death compared to female patients, whose curves appeared to be more sloping and stable. Thus, gender also influences survival rates, where female patients have a higher chance of survival during the observation period.

Figure 2c illustrates the difference *in survival* by age group. Patients ≤ 50 years of age showed a relatively high and stable survival curve, with a slight decrease over the observation time. In contrast, patients > 50 years of age experienced a faster decline *in survival* function, reflecting a higher risk of death. This suggests that older age is associated with a decreased probability of survival of kidney failure.

3.3 Log-Rank Test

The Log-Rank *test* was conducted to find out if there were significant differences in *survival* function between categories in each factor, namely disease severity, sex, and age group. The test results are presented in Table 2.

Table 2. *Log-Rank Test* Results for Each Variable

Variable	*Log-Rank*	Df	*P-Value*	Sig.	Conclusion
Severity Factors	0	1	1	0,05	Failed to Decline H_0
Gender Factors	0,2	1	0,6	0,05	Failed to Decline H_0
Age Factor	0	1	0,7	0,05	Failed to Decline H_0

Because all *p-values* were greater than the significance level of 0.05, there were no significant differences in *survival* function between categories in severity, sex, or age factors. Thus, these three factors did not have a significant effect on the risk of death during the observation period of patients with kidney failure.

4 Discussion

Based on research [4] and study [8] applying the Nelson-Aalen method, the authors estimated the cumulative *hazard* function to see the pattern of mortality risk in kidney failure patients based on disease severity, sex, and age group. The results of the cumulative *hazard* curve show that there is a difference in the pattern of increasing risk between categories. Patients with acute conditions have a faster increased risk of death than patients with chronic conditions. In addition, male patients and the 50-year-old age group also showed a steeper >*hazard* curve, indicating a higher risk in a shorter period of time.

Nevertheless, the results *Log-Rank Test* that is done based on [9] and [11] to test the difference in survival function between groups, there was no significant difference in the three factors analyzed. The entire *p-value* at the *Log-Rank Test is* greater than the significance level of 0.05, so it fails to be rejected. This means that there is no statistically significant difference in the probability of survival based on disease severity, sex, or age group. H_0

Nevertheless, these results are in line with the findings in the study [12] that uses the same data but applies an approach *Kaplan-Meier*, where no significant differences in survival probability were found based on disease severity, sex, or age group. These similarities in results suggest that although there are visual differences in risk patterns, statistically, these clinical characteristics have not shown a significant influence on the survival function *of* patients with renal failure in the observed period analyzed.

5 Conclusion

Analysis using *the Nelson-Aalen* method showed a difference in the pattern of mortality risk and survival chance in kidney failure patients based on disease severity, sex, and age. Visually, patients with acute, male, and > age 50 years appeared to have a higher risk of death during the observation period. In contrast, patients with chronic conditions, women, and 50-year-olds showed better survival tendencies.≤

However, these differences cannot be statistically ascertained through *the Log-Rank Test*. These findings emphasize the importance of further analysis with larger sample sizes or additional approaches so that the influence of each factor on the *survival* function can be more convincingly understood.

Reference

[1] Adryansyah, "Kidney Disease," [Online]. Available on: https://www.halodoc.com/kesehatan/penyakit-ginjal, accessed 24 May 2025.

[2] D. P. Kurniawati, H. Khuluq, and A. N. Ainni, "Risk Factors Affecting the Length of Hospitalization of Chronic Kidney Failure Patients," *Journal of 'Aisyiyah Medika*, vol. 10, no. 1, p. 499–509, 2025.

[3] G. S. Pongsibidang, "Risk of Hypertension, Diabetes, and Consumption of Herbal Drinks in the Incidence of Chronic Kidney Failure at Dr. Wahidin Sudirohusodo Hospital, Makassar in 2015," *Copyright © 2019 Wiyata. All Rights Reserved.*, vol. 3, no. 2, p. 162–167, 2016.

[4] D. Yulianto, H. B. Notobroto, and Widodo, "Analysis of Survival of Chronic Kidney Disease Patients with Hemodialysis at Dr. Soetomo Hospital Surabaya," *Journal of Health Management of the Dr. Soetomo Hospital Foundation*, vol. 3, no. 1, p. 96, 2017.

[5] R. Pranandari and W. Supadmi, "Risk Factors for Chronic Kidney Failure in the Hemodialysis Unit of Wates Kulon Progo Hospital," *Pharmaceutical Magazine*, vol. 11, no. 2, p. 415–418, 2015.

[6] F. Rachman, S. Wahyuningsih, and Y. Novia, "Analysis *Survival* Length of Treatment and Recovery Rate of Drug Patients at the Ibadurrahman Tenggarong Seberang Islamic Boarding School Therapy and Rehabilitation Institution," *Exponential Journal*, vol. 7, no. 1, p. 41–50, 2016.

[7] C. Khotimah, "Additive *Survival* Least Square Svm For Data Analysis *Survival*," *Thesis, Sepuluh Nopember Institute of Technology*, 2018.

[8] N. R. Dianti, B. A. P. Martadiputra, and H. S. Husain, "Curve Analysis *Nelson-Aalen* Using the Weighted Log Rank Test," *EurekaMatika Journal*, vol. 11, no. 1, p. 49–58, 2023.

[9] F. Putri, B. Martadiputra, and Lukman, "*Log-Rank Test* as A Continuation of The *Kaplan-Meier* Method in *Survival* Analysis Using R Language (Case Study: D-penicillamine Treatment on the Probability of *Survival* in Primary Biliary Cirrhosis (PBC) Patients at the Mayo Clinic)," *Journal of Mathematics, Statistics, and Computing*, vol. 21, no. 1, p. 285–306, 2024.

[10] T. Widiharih, A. Rusgiyono, Sudarno, and B. A. Saputra, "*Kaplan-Meier* and *Nelson-Aalen* Estimators for Credit Scoring," *Media Statistics*, vol. 16, no. 1, p. 37–46, 2023.

[11] S. Auliana, "Analysis of the Life Test of the Kaplan Meier Method with Log Rank Test," *Thesis, Sunan Kalijaga State Islamic University, Yogyakarta*, 2021.

[12] I. N. Divine, "Analysis *Survival* for patients with kidney failure using the Kaplan-Meier method," *Thesis, Alauddin State Islamic University Makassar*, 2022.

[13] R. Hidayat, "The Use of Methods *Kaplan-Meier* and Life Table Analysis *Survival*," *Journal of Dynamics*, vol. 7, no. 1, p. 1–8, 2016.

[14] D. I. Pratiwi, T. Wuryandari, and Sudarno, "The Use of Durability Analysis for Determination of Warranty Period and Product Price on Fluorescent Lamp Life Time Data," *Gaussian Journal*, vol. 4, no. 3, p. 463–476, 2015.

[15] A. Ikrom, Tarno, and T. Widiharih, "Survival Analysis in Patients with Pulmonary Tuberculosis using the Proportional Cox Regression Method *Hazard*," *Gaussian Journal*, vol. 13, p. 180–188, 2024.

[16] A. W. R. A. Aziz, Y. N. Nasution, and S. Wahyuningsih, "Analysis of Recurring Event Data Is Not Identical to the Cox Gap Time Model," *Journal* Exponential, vol. 8, p. 169–178, 2017.

[17] A. Suhartini, R. Rahmawati, and Suparti, "Curve Analysis *Survival* Kaplan Meier uses the log rank test," *Gaussian Journal*, vol. 7, no. 1, p. 33–42, 2018.