

Survival Analysis of Prostate Cancer Patients Using Cox Regression Model and Log-Logistic Model

Samuel Olayemi Olanrewaju¹, Ezekiel Kehinde Adeniran²

Department of Statistics, University of Abuja, Abuja, Nigeria

DOI: <https://doi.org/10.51583/IJLTEMAS.2026.150100051>

Received: 16 January 2026; Accepted: 23 January 2026; Published: 06 February 2026

ABSTRACT

Survival time analysis focuses on the time until an event occurs and is used to identify risks in survival data. This study employs Non-Parametric (Kaplan-Meier) methods to assess median survival time, Log-rank tests to compare hazard and survivor functions, Semi-Parametric (Cox Proportional Hazards), and Parametric approaches to determine the best-fitting distribution. Prostate Cancer (PC) is the second most common malignancy in men worldwide, with 1,276,106 new cases and 358,989 deaths in 2018 (Rawla, 2019). The incidence and mortality of prostate cancer increase with age, with the average diagnosis age being 66 years. African-American men have higher incidence rates (158.3 new cases per 100,000 men) and nearly double the mortality rate compared to White men (Capece et al., 2020). This study found that both the age of patients and the year of admission were consistently significant factors. The Log-logistic model was identified as the bestfitting model with an AIC value of 302.7047, compared to the Cox Regression model's AIC value of 434.0985.

INTRODUCTION

Prostate cancer is the second most frequent malignancy in men worldwide, with 1,276,106 new cases and 358,989 deaths in 2018 (Rawla, 2019). The incidence and mortality rates of prostate cancer correlate with increasing age, with the average diagnosis age being 66 years. African-American men have higher incidence rates compared to White men, with 158.3 new cases diagnosed per 100,000 men and nearly double the mortality rate (Capece et al., 2020). This disparity may be due to social, environmental, and genetic factors. By 2040, 2,293,818 new cases are estimated, with a small variation in mortality (Tosoian et al., 2015).

Prostate cancer often presents with no symptoms and progresses slowly, with common complaints including nocturia and difficulty urinating. Male-specific antigens (PSA > 4 ng/mL) are used for identifying prostate malignancies, though a biopsy is required for confirmation. Dietary and exercise factors significantly influence the onset and spread of prostate cancer (Capurso & Vendemiale, 2017). The disease continues to pose a significant public health challenge in Nigeria, where its burden appears increasingly substantial due to both rising incidence and late clinical presentation. National evidence suggests that prostate cancer incidence and prevalence have been increasing over time, with a wide range of hospital-based prevalence estimates and high mortality observed within three years of diagnosis, reflecting advanced disease at presentation in many Nigerian settings (Epidemiology of prostate cancer in Nigeria, 2024).

In community and clinical settings, studies consistently reveal low levels of knowledge about prostate cancer and very limited screening uptake. For example, in Sokoto State only 5% of surveyed men were aware of prostate cancer and just 1.3% knew about available screening tests, with no participant reporting prior screening, largely due to lack of awareness (Awosan et al., 2018). Similarly, in Ido-Ekiti only 18.2% of men aged 40 and above reported ever having a prostate cancer screening test, despite moderate awareness, underscoring systemic gaps in screening access (Adewoye et al., 2023). Further research in Bayelsa communities reported low knowledge about prostate cancer symptoms, prevention, and management, highlighting the broad need for education and early detection programs (Awareness of Prostate Cancer..., 2025).

Collectively, this body of Nigerian evidence highlights critical gaps in awareness, screening, and early detection, and supports the need for rigorous survival analyses that reflect the real-world experiences of Nigerian prostate cancer patients. These findings provide essential context for survival studies, as late presentation, often due to low screening uptake and poor knowledge, is intrinsically linked with poorer survival outcomes in low-resource settings.

Statement of Problem

Prostate cancer (PC) is the fourth major cancer globally, accounting for 1.3 million cases (7.1% of the overall cancer incidence) (Shafique et al., 2012). It is the second most frequently diagnosed cancer disease in men, with varying incidence rates globally (Sebastiao & Peter, 2018). In southwestern Nigeria, PC is prevalent among men aged 46-99 years, with a peak incidence in those ≥ 70 years old (Van Wijk & Simonsson, 2022).

Risk factors for PC include both non-modifiable factors (age, race, family history) and modifiable factors (alcohol consumption, obesity, smoking, sedentary lifestyle, prostatitis history, high cholesterol) (Balogun et al., 2020). This study examines the risk factors associated with prostate cancer in Nigeria and models the survival pattern of prostate cancer patients.

Aim and Objectives

This study aims to compare Cox regression and parametric models for the survival of prostate cancer patients. The specific research objectives are to:

- i. Estimate the survival time with respect to variables of interest.
- ii. Compare hazard and survivor functions of different variables of interest.
- iii. Fit the Cox proportional hazard model along with different parametric models to the prostate cancer data.
- iv. Choose the best-fitted model for the prostate cancer data.

RESEARCH METHODOLOGY

Source of Data

The data for this study are secondary data obtained from the records of six hundred and sixty-one (661) registered prostate cancer patients at the University of Ilorin Teaching Hospital. The data collected is based on the length of stay, age, gender, and outcomes over a twelve-year period (2011-2022). To facilitate computation and make the best use of the statistical tools applied in this research, the covariates were categorized as follows:

Table 3.0: Categories of Covariates

Covariates	Categories
Gender	Male and Female
Age (Years)	5-19, 20-39, 40-59, 60-79 and 80-99
Year of Admission	2011-2022
Outcome	Dead or Alive

RESEARCH METHODOLOGY

Kaplan Meier Estimator

The technique used to estimate the survivor function when censoring is present is called the Kaplan-Meier or the product-limit estimator of the survivor function (or survival probability). Equation 1 below displays the Kaplan-Meier estimator of survival at time t . In this case, d_j is the number of failures at time t_j , r_j is the number of people at risk at time t_j , and $t_j, j = 1, 2, \dots, n$ is the whole set of failure times recorded (with t^+ the greatest failure time).

$$S(t) = \prod_{j: t_j \leq t} \left(\frac{r_j - d_j}{r_j} \right), \text{ for all } t \leq t^+ \quad (1)$$

It is simply the empirical probability of surviving past certain times in the sample (taking into account censoring). When there is inappropriate censoring, the Kaplan Meier method is not appropriate. The general formula for a KM survival probability at failure time $t(j)$ is shown below;

$$S(t(j)) = S(t(j-1)) \times \Pr[T > t(j) | T \geq t(j)] \quad (2)$$

This formula gives the probability of surviving past the previous failure time $t(j-1)$, multiplied by the conditional probability of surviving past time $t(j)$, given survival to at least time $t(j)$. The above KM formula can also be expressed as a product limit if the survival probability $S(t(j-1))$ is substituted, the product of all fractions that estimate the conditional probabilities for failure times $t(j-1)$ and earlier. This is expressed as;

$$S(t(j-1)) = \prod \Pr[T > t(i) | T \geq t(i)] \quad (3)$$

$$S(t) = \prod_{j: t_j \leq t} \left(\frac{r_j - d_j}{r_j} \right), \text{ for } 0 \leq t \leq t^+ \quad (4)$$

Equation 4 is the empirical probability of surviving past certain times in the sample (taking into account censoring).

Log Rank Test

The Log rank test also known as Mantel- Haenszel test, is a large sample chi-square test that uses as its test criterion a statistic that provides an overall comparison of the KM curves being compared (Sebastiao and Peter, 2018). It is applicable to data where there is progressive censoring and gives equal weight to early and late failures. The test statistic can be expressed as;

$$\chi^2 = \sum \frac{(O_i - E_i)^2}{E_i}, i = 1, 2, \dots, \chi^2_{1, df} \quad (5)$$

$$Var(O_i - E_i)$$

The log-logistic model is discussed below because it is appropriate for cancer survival data where the risk of death may vary over time rather than remain constant. In prostate cancer, mortality risk often increases after diagnosis and treatment before stabilizing among longer-term survivors, a pattern the log-logistic model can capture through its flexible hazard function. Unlike models that assume monotonic hazards, log-logistic distribution allows for this non-monotonic behaviour. Its suitability in this study is further supported by lower AIC and BIC values compared with the Cox proportional hazards model, indicating a better fit to the data.

Log-Logistic Model

The log-logistic (LL) distribution (branded as the Fisk distribution in economics) possesses a rather simple functional form. The LL distribution is among the class of survival time parametric models where the hazard rate initially increases and then decreases and at times can be hump-shaped (Adelian et al., 2015). The hazard rate function of LL model is given below:

$$h(x, \alpha, \lambda) = \frac{\alpha}{x^\alpha} \quad (6)$$

$$x[l+(x/\lambda)]$$

The log-logistic (LL) model is expressed mathematically as;

$$\lambda^{-1/p} = \exp(\alpha_0 + \alpha_1 X_1 + \alpha_2 X_2 + \dots + \alpha_p X_p) \quad (7)$$

Where the mean and the variance of the log-logistic model is given as:

$$\text{Mean} = \frac{\pi/\beta}{\sin(\pi/\beta)}, \text{ if } \beta > 1, \text{ else undefined.} \quad (8)$$

$\sin(\pi/\beta)$

Cox-proportional Hazard Model

The Cox-proportional hazards model possess the property that different individuals have hazard functions that are proportional i.e. $[h(t|x_1)/h(t|x_2)]$, the ratio of the hazard functions for two individuals with prognostic factors or covariates $x_1=(x_{11}, x_{21}, \dots, x_{p1})'$, and $x_2=(x_{12}, x_{22}, \dots, x_{p2})$ is a constant (does not vary with time t). This means that the ratio of the risk of dying of two individuals is the same no matter how long they survive. The log hazard's linear-like model must also be specified for this model. The exponential distribution can be used to parameterize a parametric model in the following way:

$$\log h_i(t) = \alpha + \beta_1 x_{1i} + \beta_2 x_{2i} + \dots + \beta_k x_{ki} \quad (9)$$

In this case, the constant α represents the log-baseline since $h_i(t) = \alpha$, when all the x 's are zero. The Cox proportional hazards model is a semi-parametric model where the baseline hazard $\alpha(t)$ is allowed to vary with time.

RESULTS AND DISCUSSION

This chapter discusses the analysis of data collected on prostate cancer patients using three methods: Nonparametric methods (including Kaplan-Meier and Log-Rank Statistic), Semi-parametric (Cox Proportional Hazard), and a parametric model (Log-normal). The results are present

Kaplan-Meier Survival Curves

The figures below show the general Kaplan-Meier survival curve for all prostate cancer patients as well as for each covariate considered in this study.

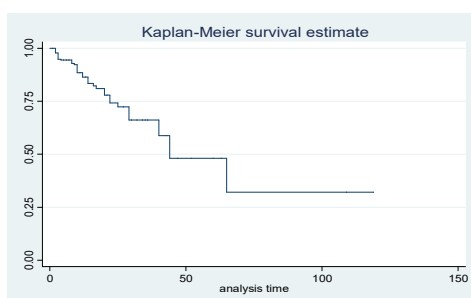


Figure 4.1: General survival curve for prostate cancer patients.

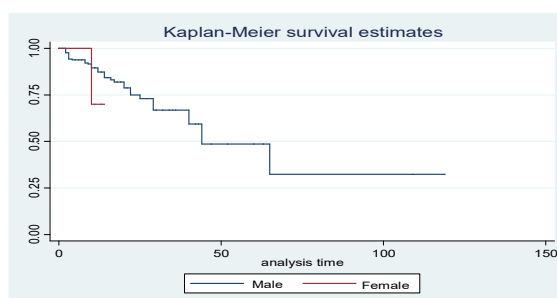


Figure 4.2: Survival curve on sex of prostate cancer patients with log-rank p-value = 0.6446.

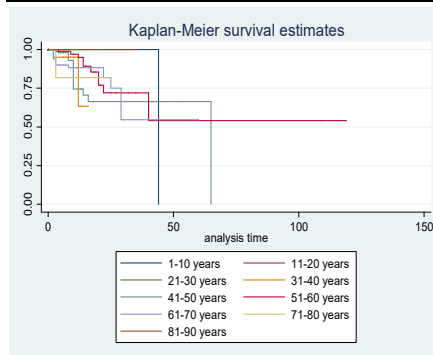


Figure 4.3: Survival curve on age prostate cancer patients distribution of prostate with log-rank p-value = 0.0179.
rank p-value = 0.0372.

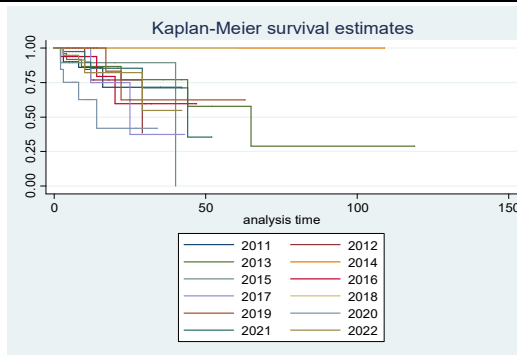


Figure 4.4: Survival curve on year cancer patients with log-rank p-value = 0.0372.

Figure 4.1 shows the survivorship function based on the total number of patients' data collected with the number at risk at different time points. Figure 4.2 reveals the survivorship function based on the sex of prostate cancer patients. While Figure 4.3 indicates the survivorship function based on age categories of prostate cancer patients and Figure 4.4 shows the survivorship function based on the year of admission of prostate cancer patients with the number at risk at different time points.

Log-Rank Test for Equality of Survivor Functions

4.2.1 Log-rank Statistic

$$\sum_{i=1}^n \frac{O_i - E_i}{E_i} \sim \chi^2_{(df)}$$

H_0 : Survival curves for covariates are the same

V_s

H_1 : Survival curves for covariates are not the same

Decision rule: Reject H_0 if p-value < α -level (0.05). Otherwise, do not reject H_0 .

Table 4.1: Log-rank Test for the Sex of the Prostate Cancer (PC) Patients

SEX	EVENT OBSERVED	EVENT EXPECTED
Male	43	43.67
Female	3	2.33
Total	46	46.00

$$\chi^2_{(1)} = 0.21$$

$$Pr > \chi^2 = 0.0410$$

Decision: Since the p-value = 0.0410 < $\alpha = 0.05$. The null hypothesis (H_0) is therefore rejected.

CONCLUSION

Since the null hypothesis has been rejected. It can therefore be concluded that patients' sex has significant different K-M survival curves. This means that sex of prostate cancer patients does affect their survival.

Table 4.2: Log-rank Test for the Different Age Categories of Prostate Cancer Patients

AGE	EVENT OBSERVED	EVENT EXPECTED
1-10 years	2	1.45
11-20 years	0	0.52
21-30 years	0	0.06
31-40 years	3	1.99
41-50 years	12	10.08
51-60 years	11	16.50
61-70 years	14	12.87
71-80 years	4	2.29
81-90 years	0	0.25
Total	46	46.00

$$\chi^2_{(8)} = 5.35$$

$$Pr > \chi^2 = 0.0179$$

Decision: Since the p-value= 0.0179 < $\alpha=0.05$. The null hypothesis (H_0) is therefore rejected.

CONCLUSION

Since the null hypothesis has been rejected. It can therefore be concluded that different categories of the patients' age have significant different K-M survival curves. This means that patients' age of prostate cancer patients does affect their survival.

Table 4.3: Log-rank Test for the Years of Admission of Prostate Cancer Patients

YEARS OF ADMISSION	EVENT OBSERVED	EVENT EXPECTED
2011	5	4.27
2012	4	3.02
2013	6	7.19
2014	0	2.20
2015	4	3.47
2016	6	5.62
2017	2	1.49
2018	0	1.81
2019	2	3.36
2020	5	1.49

2021	6	7.16
2022	6	4.91
Total	46	46.00

$$\chi^2_{(11)} = 15.06$$

$$Pr > \chi^2 = 0.0372$$

Decision: Since the p-value = 0.0372 < $\alpha = 0.05$. The null hypothesis (H_0) is therefore rejected.

CONCLUSION

Since the null hypothesis has been rejected. It can therefore be concluded that different years of admission have significant different K-M survival curves. This means that patients' year of admission do affect patients' survival rate.

Test Of Proportional Hazards Assumption

H_0 : The Proportional Hazard assumption is not violated

Vs

H_1 : The Proportional Hazard assumption is violated

Time: log(t)

Table 4.4: Proportional Hazard Assumption

Covariates	Rho	chi2	df	prob>chi2
SEX	-	-	-	-
Female	0.05524	0.15	1	0.0248
AGE Categories	-	-	1	-
11-20 years	-	-	1	-
21-30 years	-	-	1	-
31-40 years	-	-	1	-
41-50 years	-0.17575	1.23	1	0.2671
51-60 years	-0.32579	3.52	1	0.0607
61-70 years	-0.24246	1.8	1	0.1800
71-80 years	-0.27858	2.61	1	0.1059
81-90 years	-0.30277	3.3	1	0.0692
Years of Admission	-	-	-	-
2012	0.11455	0.61	1	0.4366
2013	-0.07376	0.28	1	0.5965

2014	-	-	1	-
2015	0.18845	1.57	1	0.2108
2016	-0.02635	0.03	1	0.8570
2017	0.07262	0.26	1	0.6085
2018	-	-	1	-
2019	0.09379	0.37	1	0.5413
2020	-0.03495	0.06	1	0.8046
2021	0.02286	0.02	1	0.8757
2022	0.05076	0.13	1	0.7220
Global Test	-	9.95	15	0.8227

The test Table 4.4 do not suggest violation of the PH assumption for any of the covariates with their p-values greater than 0.05. thus, the cox model will be fitted to the prostate cancer data.

COX PROPORTIONAL HAZARD MODEL

No. of subjects = 309

Number of observations = 309

No. of failures = 46

LR χ^2 (15) = 20.17

Time at risk = 3902

Prob > χ^2 = 0.1656

Log likelihood = -202.04923

Table 4.5 Cox Regression Model

Covariates	Coefficient (β)	Hazard Ratio	Std. Err.	z	p> z 	[95% Conf. Interval]
SEX	-	-	-	-	-	-
Female	0.4033	1.4967	0.6542	0.62	0.038	-0.8789 1.6855
AGE Categories	-		-	-	-	-
11-20 years	-43.2297	1.68e-19	-	-	-	-
21-30 years	-42.4671	3.60e-19	-	-	-	-
31-40 years	-0.1435	0.8663	1.0135	-0.14	0.887	-2.1299 1.8429

41-50 years	-0.4262	0.6529	0.8296	-0.51	0.607	-2.0521 1.1997
51-60 years	-0.9063	0.4040	0.8600	-1.05	0.292	-2.5919 0.7793
61-70 years	-0.5354	0.5854	0.8677	-0.62	0.537	-2.2361 1.1652
71-80 years	0.0447	1.0458	0.9892	0.05	0.964	-1.8941 1.9835
81-90 years	-0.6815	0.5058	-	-	-	-
Years of Admission	-	-	-	-	-	-
2012	0.1771	1.1937	0.6981	0.25	0.800	-1.1912 1.5453
2013	-0.4286	0.6514	0.6790	-0.63	0.528	-1.7595 0.9022
2014	-45.2978	2.13e-20	-	-	-	-
2015	0.0996	1.1048	0.6852	0.15	0.884	-1.2435 1.4427
2016	0.0043	1.0043	0.6332	0.01	0.995	-1.2367 1.2453
2017	0.2853	1.3301	0.8848	0.32	0.747	-1.4489 2.0195
2018	-45.3221	2.07e-20	-	-	-	-
2019	-0.3839	0.6812	0.8616	-0.45	0.656	-2.0726 1.3046
2020	1.2129	3.3634	0.6729	1.80	0.071	-0.1059 2.5318
2021	-0.3268	0.7212	0.6423	-0.51	0.611	-1.5857 0.9321
2022	0.0628	1.0648	0.6413	0.10	0.922	-1.1941 1.3196

The Estimated model: $\hat{h}(t, X) = h_0(t)e^{\beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3}$. The Cox's Proportional Hazard Model was employed to determine the hazard ratio of the groups of covariates.

Key Observations:

The hazard ratio for male relative to female is 1.4967, indicating that male patients have a longer survival time than female patients, with female patients having a 1.4967 times higher risk of dying from prostate cancer.

The proportional hazards assumption was formally assessed using Schoenfeld residuals. The test results indicated that none of the covariates violated the proportional hazards assumption, and the global test was also not statistically significant, suggesting that the hazard ratios remained constant over time. This confirms that the Cox proportional hazards model was appropriate for analyzing the survival data in this study and that the estimated effects of the covariates can be reliably interpreted. The absence of significant violations further strengthens the validity of the Cox model results presented.

Table 4.6: Akaike information criterion and Bayesian information criterion						
Model	Observation	11 (null)	11 (model)	df	AIC	BIC
	309	-212.133	-202.0492	15	434.0985	490.0986

Table 4.6 shows the Akaike information criterion and Bayesian information criterion generated for the cox regression model which will later be used to compare the other model fitted to the prostate cancer data. Thus, the best model would be chosen among the fitted models.

Parametric Method (Log-logistic Model)

No. of subjects = 309 Number of observations = 309

No. of failures = 46 $LR^2(15) = 0.19$

Time at risk = 3902 $Prob > \chi^2 = 0.9796$

Log likelihood = -146.35234

Table 4.7: Log-logistics Model

Covariates	Coefficient	Std. Err.	z	p> z	[95% Conf. Interval]	
SEX	-0.1518	0.4908	-0.31	0.001	-1.1137	0.8101
AGE	-0.0154	0.0876	-0.18	0.050	-0.1871	0.1563
YEARS OF ADMISSION	-0.0079	0.0344	-0.23	0.017	-0.0754	0.0595
_Cons	4.2609	0.8413	5.06	0.001	2.6119	5.9099
/ln_gam	-0.2845	0.1136	2.50	0.012	-0.5073	-0.0618
/gamma	0.7524	0.0855			0.6021	0.9400

The Log-logistic model fitted to the data is as follows;

$$\frac{I}{I/p} (\alpha_0 + \alpha_1 X_1 + \alpha_2 X_2 + \dots + \alpha_p X_p) = \exp \left(\frac{\lambda}{\lambda} \right)$$

Therefore, the fitted model is given below;

$$\frac{I}{I/p} = \exp \left(4.2609 - 0.1518 \text{Sex} - 0.0154 \text{Age} - 0.0079 \text{Year} \right)$$

λ

Also, the model fitted for the significant covariates are as follows;

$$ll/p \quad \text{---} \quad 4.2609 - 0.1518_{sex} - 0.0154_{Age} - 0.0079_{Year}$$

$$= \exp \lambda$$

Table 4.8:	Akaike information criterion and Bayesian information criterion					
Model	Observation	11 (null)	11 (model)	df	AIC	BIC
	309	- 146.4462	-146.3523	5	302.7047	321.3714

Table 4.8 reveals the Akaike information criterion and Bayesian information criterion generated for the Loglogistic model which will later be used to compare Cox model fitted to the prostate cancer data. Therefore, the model with the least AIC would be selected as the best model fitted to the prostate cancer data.

Table 4.9:	MODEL SELECTION			
Model	Number of Parameters (p)	Log Likelihood	AIC	BIC
Cox Regression	3	-202.0492	434.0985	490.0986
Loglogistic	4	-146.3523	302.7047	321.3714

For the best model to be chosen, the AIC values for the fitted models are compared. Thus, the best fitted model selected for the prostate cancer data is Loglogistic model having the least AIC and BIC value respectively.

SUMMARY OF FINDINGS

This study aimed to identify the factors that affect the survival level of prostate cancer (PC) patients. Key findings include:

- **Kaplan-Meier Survival Estimates:** Sex, age, and year of admission significantly affected the survival rate of PC patients.
- **Log-Rank Test:** The survival experience of patients differed significantly based on sex, age, and year of admission at a 0.05 significance level.
- **Cox Regression and Loglogistic Models:** Both models identified sex, age, and year of admission as significant covariates. These factors consistently contributed to the models.
- **Best Model:** The Loglogistic model was the best fit for the prostate cancer data, with the least AIC value of 302.7047.

CONCLUSION

Based on the analysis, the following conclusions were made:

- Sex, age, and year of admission significantly affect the survival rate of PC patients at a 0.05 level of significance.
- Both Cox and Loglogistic models identified sex, age, and year of admission as significant covariates.

- The covariates consistently significant in both models were sex, age, and year of admission.

RECOMMENDATIONS

Based on this research, the following recommendations are made:

- The year of admission of prostate cancer (PC) patients is crucial. People with PC should seek hospital treatment as early as possible for better outcomes.
- Parametric models are recommended for analyzing PC data. The Loglogistic model is the best fit for PC data irrespective of sample sizes.

Limitation of the Study

Although this study examined key demographic and temporal covariates such as age, sex, and year of admission, other clinically relevant factors that may influence prostate cancer survival were not included in the analysis. These include treatment modality, disease stage at diagnosis, comorbid conditions, and socio-economic status, which have been shown in previous studies to affect cancer outcomes. The absence of these variables in the hospital records limited the ability to fully adjust for potential confounding effects, and their omission should be considered when interpreting the findings.

In the Nigerian healthcare context, variations in access to diagnostic facilities, treatment availability, and followup care may further contribute to differences in survival outcomes. Patients are often present at advanced stages due to low screening uptake and delayed healthcare-seeking behaviour, factors that are closely linked to both disease progression and treatment response. Future studies incorporating clinical staging information, treatment data, and indicators of socio-economic status would provide a more comprehensive understanding of prostate cancer survival and support more targeted interventions in Nigeria.

Suggestions for Further Study

This study focused on the survival rate of prostate cancer using data from the University of Ilorin Teaching Hospital (UIH). Suggestions for future research include:

- Employ different parametric models other than the Loglogistic model to compare the survival levels of prostate cancer.
- Compare other semi-parametric models aside from the Cox Regression Model for analyzing the survival level of prostate cancer patients.

REFERENCES

1. Adewoye, K. R., Aremu, S. K., Adegbiyi, W. A., & Achebe, C. C. (2023). Awareness, knowledge, and factors that influenced the uptake of screening tests for prostate cancer among men aged 40 and older in Ido-Ekiti, Ekiti State, Nigeria. *Journal of Public Health in Africa*, 14(4), Article 1634. <https://doi.org/10.4081/jphia.2023.2134>
2. Awareness of prostate cancer among residents in Amasomma and Ogubiri communities in the Niger Delta Region, Bayelsa State, Nigeria. (2025). *Cancer Therapy & Oncology International Journal*. <https://juniperpublishers.com/ctoj/CTOIJ.MS.ID.556202.php>
3. Balogun, O. S., Gao, X. Z., Jolayemi, E. T., & Olaleye, S. A. (2020). Generalized cure rate model for infectious diseases with possible co-infections. *Plos one*, 15(9), e0239003.
4. Capece, M., Creta, M., Calogero, A., La Rocca, R., Napolitano, L., Barone, B., ... & Longo, N. (2020). Does physical activity regulate prostate carcinogenesis and prostate cancer outcomes? A narrative review. *International journal of environmental research and public health*, 17(4), 1441.

5. Capurso, C., & Vendemiale, G. (2017). The Mediterranean diet reduces the risk and mortality of the prostate cancer: A narrative review. *Frontiers in Nutrition*, 4, 38.
6. Epidemiology of prostate cancer in Nigeria: a mixed methods systematic review. (2024). *BMC Cancer*, 24, Article 2511. <https://pubmed.ncbi.nlm.nih.gov/39306811/>
7. Sebastiao, Y. V., & Peter, S. D. S. (2018, December). An overview of commonly used statistical methods in clinical research. In *Seminars in pediatric surgery* (Vol. 27, No. 6, pp. 367-374). WB Saunders.
8. Shafique, K., McLoone, P., Qureshi, K., Leung, H., Hart, C., & Morrison, D. S. (2012). Cholesterol and the risk of grade-specific prostate cancer incidence: evidence from two large prospective cohort studies with up to 37 years' follow up. *BMC cancer*, 12(1), 1-8.
9. Tosoian, J. J., Mamawala, M., Epstein, J. I., Landis, P., Wolf, S., Trock, B. J., & Carter, H. B. (2015). Intermediate and longer-term outcomes from a prospective active-surveillance program for favorable-risk prostate cancer. *Journal of Clinical Oncology*, 33(30), 3379.
10. Van Wijk, R. C., & Simonsson, U. S. (2022). Finding the right hazard function for time-to-event modeling: A tutorial and Shiny application. *CPT: Pharmacometrics & Systems Pharmacology*, 11(8), 991-1001.