




SURVIVAL QUANTILE REGRESSION ANALYSIS OF COVID-19 IN RIVERS STATE, NIGERIA

 **Maureen T. Nwakuya**^{1*}

 **Omosivie Maduka**²

¹Department of Mathematics/Statistics, University of Port Harcourt, Port Harcourt Rivers State, Nigeria.

Email: maureen.nwakuya@uniport.edu.ng Tel: +2348033167003

²Department of Preventive and Social Medicine, College of Health Sciences, University of Port Harcourt, Port Harcourt Rivers State, Nigeria.

Email: omosivie.maduka@uniport.edu.ng Tel: +2348033298096



(+ Corresponding author)

ABSTRACT

Article History

Received: 3 February 2021

Revised: 17 February 2021

Accepted: 26 February 2021

Published: 4 March 2021

Keywords

Survival quantile regression
Cox proportional Hazard model
Accelerated failure time model
Log-rank test
Survival function and censoring.

Coronavirus 2019 (Covid-19) cases in Rivers State, Nigeria are on the increase day by day. It became imperative to investigate the survival rate of covid-19 patients in this state. The survival quantile regression was applied assuming right censoring to estimate the effect of age, sex, fever, anosmia, comorbidity, and cough on the survival time of patients. The results show that on admission into the hospital the survival time of the patients depended on the age and the presence of anosmia, comorbidity, and fever. By the mid survival period only anosmia and fever were seen to be significant but at the 75th quantile comorbidity was also seen to be significant along with fever and anosmia. The result also shows that having fever is associated with longer stay in the hospital based on the size of the effect at different quantiles. We also noticed that though the effect of anosmia and comorbidity were significant at the 25th and 75th quantile the sizes of the effects were minimal, but comorbidity was seen to have a bigger effect than anosmia. Comparing the survival time of groups, the results showed that males and females have the same survival time and patients with and without comorbidity equally have the same survival time. Patients without fever, anosmia and cough had a shorter survival time than those that had fever, anosmia, and cough. We then concluded that fever, comorbidity, and anosmia are the major factors that affect the survival time of covid-19 patients in Rivers State, Nigeria.

Contribution/Originality: This study is one of the very few studies that have investigated the effect of different covariates at different points on the distribution of the survival time of Covid-19 patients in Rivers State, Nigeria.

1. INTRODUCTION

Coronavirus 2019 (Covid-19) cases in Rivers State, Nigeria are on the increase day by day. It became imperative to investigate the survival rate of Covid-19 patients. Survival analysis is widely used in modelling and analyzing time-to-event data. Survival analysis is said to be a collection of statistical procedures for data analysis for which the outcome variable of interest is time until an event occurs, often referred to as a failure time, survival time, or event time. Survival time refers to a variable which measures the time from a particular starting time (e.g., time diagnosed of sickness) to a particular endpoint of interest, say time discharged/case resolved. The special feature about survival analysis is that in real life situations some of the event times cannot be observed due to a variety of reasons, like drop out or death before the issue is resolved. When there is sure missing information, it is termed censoring. Most researchers are often interested in estimating the survival function (or distributions of survival times), comparing survival among groups and finding the relationship between covariates and survival times. In the regression model approach for survival analysis the Cox proportional hazards model and accelerated failure time model are widely used. The introduction of quantile regression has attracted considerable interest in survival analysis. It offers a more valuable complement to the Cox proportional hazards model Cox [1] and the accelerated

failure time model Buckley and James [2] in survival analysis. Quantile regression allows the covariate effects to vary at different tails of the survival time distribution. Such important heterogeneity in the population may be neglected by using the Cox model and the accelerated failure time model, hence making survival quantile regression the best alternative.

The Covid-19 virus which was first confirmed in China towards the later part of 2019 has been on the rampage across the globe. The projections of its havoc in Africa was met with a lesser impact compared to the rest of the world. Report from the World Health Organization (WHO) as at December 31st 2020, has it that there are 41,329,493 confirmed cases in America and 30,488,064 in Europe while Africa is tailing with a total number of 2,313,130 confirmed cases. Some researchers have proposed that this could be as a result of weather conditions, etc, being that as it may, we decided to investigate the effect of some factors on the survival time of covid-19 patients. Using a covid-19 dataset this paper provides a practical guide to using quantile regression in survival analysis of covid-19 patients. Thus this paper therefore focuses on application of survival quantile model to a covid-19 data. The rest of the paper is as follows; the first section presents the basic notations in survival analysis, followed by explanations of censoring, Cox proportional hazards model and accelerated failure time model, following this section is; survival quantile regression model, methodology, results and conclusions.

2. BASIC NOTATIONS IN SURVIVAL ANALYSIS

In survival analysis the function that gives the probability of the survival time occurring at exactly time t is the density function in Equation 1 ;

$$f(t) = \lim_{\Delta t \rightarrow 0} \frac{P(t \leq T < t + \Delta t)}{\Delta t} = \frac{dF(t)}{dt} = -\frac{dS(t)}{dt} \tag{1}$$

The function that gives the probability that the survival time occurring before or exactly at time t is the cumulative distribution function $F(t)$, given in Equation 2;

$$F(t) = P(T \leq t) = \int_0^t f(u) du \tag{2}$$

While the survival function $S(t)$ given in Equation 3 gives the probability that a person survives longer than some specified time t : that is, $S(t)$ gives the probability that the random variable T exceeds the specified time t .

$$S(t) = P(T > t) = 1 - F(t) = 1 - \int_0^t f(u) du$$

$$P(T > t) = \int_t^{+\infty} f(u) du = e^{-\int_0^t h(u) du} \tag{3}$$

Some important characteristics of survival function are:

- It is non-increasing; that is, it heads downward as t increases.
 $t = 0, S(t) = S(0) = 1$
- At time $t = 0$; that is, at the start of the study, since no one has gotten the event yet, the probability of surviving past time zero is one.
- At time $t = \inf$, $S(t) = S(\inf) = 0$
- At time $t \rightarrow \infty$, that is, theoretically, if the study period increased without limit, eventually nobody would survive, so the survival curve must eventually fall to zero.

$$h(t) \qquad T \geq t$$

The hazard function $h(t)$ given in Equation 4 is the survival rate at time t given the survival time $T \geq t$;

$$h(t) = \lim_{\Delta t \rightarrow 0} \frac{P[t \leq T < t + \Delta t | T \geq t]}{\Delta t} \tag{4}$$

Also the hazard function can be expressed as a function of the probability density function and the survival function;

$$P[t \leq T < t + \Delta t | T \geq t] = \frac{P[t \leq T < t + \Delta t, T \geq t]}{P[T \geq t]} = \frac{f(t)}{S(t)}$$

where $P(\cdot)$ is probability.

The hazard function $h(t)$ has the following characteristics:

- It is always nonnegative, that is, equal to or greater than zero.
- It has no upper bound.

3. CENSORING

When event times cannot be fully observed due to a variety of reasons, data are subject to censoring which makes statistical estimation and inference for survival quantile regression more involved. Censoring is of different types, including; Right censoring, left censoring and interval censoring and censoring can be classified into fixed and random censoring. Any procedure which completely ignores censoring may give highly biased estimates Koenker [3]. Powell [4] first studied censored quantile regression with fixed censoring. Ying, et al. [5] proposed a semiparametric median regression model for random censoring. Despite the simplicity of the method in Ying, et al. [5] this procedure requires the unconditional independence of the survival time and censoring time. Relaxing the independence condition Portnoy [6] developed a novel estimating approach motivated by the classical Kaplan–Meier estimator in the one sample analysis. He provided a recursive estimator for right censored data, which only requires conditional independence between the survival time and censoring time given the covariates. Peng and Huang [7] also proposed a quantile regression for right censoring motivated by Nelson–Aalen estimator of the cumulative hazard function in univariate cases.

4. COX PROPORTIONAL HAZARDS MODEL

The Cox proportional hazards (CPH) model is one of the widely used models in survival analysis. It evaluates simultaneously the effect of the covariates on survival. Cox proportional hazards models links the hazard (the instantaneous rate of failure) to the covariates. In other words, it allows us to examine how specified variables influence the rate of a particular event happening at a particular point in time. This rate is commonly referred to as the hazard rate. The Cox model is expressed by the *hazard function* denoted by $h(t)$ and it is usually given in Equation 5 as;

$$h(t) = h_0(t) e^{X_i' \beta} \tag{5}$$

- t represents the survival time.
- $h(t)$ is the hazard function .
- β is a vector of coefficients that measure the impact (i.e., the effect size) of covariates.
- $h_0(t)$ is the unspecified baseline hazard function.

The model presumes that the conditional hazard of the event being modeled is the product of a baseline hazard $h_0(t)$ $e^{X_i' \beta}$

(that is common to all individuals) and the multiplicative effect of the covariates. The quantities are called the hazard ratios (HR), a hazard ratio greater than one, indicates that as the value of the covariate increases, the event hazard increases and thus the length of survival decreases.

The integrated or cumulative baseline hazard is defined in Equation 6 as;

$$H_0(t) = \int_0^t h_0(u) du \tag{6}$$

and it can be related to the conditional survival function $S(t/x)$ in Equation 7 as;

$$\log(-\log(S(t/x))) = \log H_0(t) - X_i' \beta, \tag{7}$$

where, $S(t/x)$ denotes the conditional survival function. Now evaluating both sides at the random event time T , i.e evaluating at $t = T_i$, we have the model;

$$\log H_0(t) = X_i' \beta + u_i \tag{8}$$

Where $u_i = \log(-\log S(T_i/x_i))$, since $S(T_i/x_i)$ is a uniformly distributed random variable, it follows that u_i are independent identically distributed with distribution function in Equation 9

$$F_0(u) = 1 - e^{-e^u} \tag{9}$$

5. ACCELERATED FAILURE TIME MODEL

The accelerated failure time (AFT) model provides an alternative approach to the Cox proportional hazard model. The AFT models the linear relationship between log-transformation of survival times and covariates. More specifically AFT is of the form;

$$\log(T) = X_i' \beta + \epsilon_i \tag{10}$$

With the distribution of ϵ_i specified by some parameters. The AFT directly links log of the response that is the survival time (T) to the covariates. [Koenker and Geling \[8\]](#) pointed out that the CPH and AFT are limited and that they can be treated as a special case in the quantile regression model where some appropriate transformation of the survival time T is applied to make the errors independent and identically distributed.

6. SURVIVAL QUANTILE REGRESSION MODEL

Since the inception of Quantile regression in 1978, it has emerged as a powerful and natural approach to model the heterogeneous effects of covariates for a non-homogeneous population. Unlike the Cox proportional hazards model and accelerated failure time model, quantile regression models the quantile of survival time and links it to the covariates. It is robust to outliers in the data and compared to CPH and AFT models, quantile regression relaxes the proportional hazards assumption and links the whole distribution of an outcome to the covariates of interest. Quantile regression is by far a natural choice as it helps decipher the various roles each covariate plays on different quantile levels of survival. A right censoring is considered because in this research some patients died before the covid-19 is resolved/discharged and thus the exact discharge time is unobserved. Thus this work shows the usefulness of the survival quantile regression formulation for right censored outcome and transformation models, generally.

The most basic quantile regression survival model was introduced by [Koenker and Geling \[8\]](#). Given any $\tau \in (0,1)$, the τ -quantile is a value at or below which a τ -fraction of the data lies. When the quantile is defined based on the distribution of T alone, without considering covariates, X , it is called the marginal quantile or unconditional quantile. When we consider the quantile of T [or $\log(T)$] within subgroups defined by X , for example, we refer to the τ -th conditional quantile of T given X , as denoted by $Q_{\tau|X}(\tau|X)$. This Quantile regression models the relationship between the event time and the covariates using the quantile function in [Equation 11](#);

$$Q_{\tau|X}(\tau|X) = X' \beta(\tau) \tag{11}$$

where X is a vector of covariates, $\beta(\tau)$ is to the effect of X on the τ th quantile. This model in [Equation 11](#) allows the covariate effect to change with τ . The regression quantile estimator $\hat{\beta}(\tau)$ when there is no censoring, is the minimizer of the objective function over beta, given in [Equation 12](#) as;

$$\hat{\beta}(\tau) = \underset{\beta}{\operatorname{argmin}} \sum_{i=1}^n \rho_{\tau}(T_i - x_i' \beta) \tag{12}$$

Where the “check function” $\rho_{\tau}(u) = u(\tau - I(u < 0))$.

When the responses are subject to censoring, we observe $\{T_i^*, \Delta_i, X_i\}$. In a right censoring situation, $T_i^* = \min(T_i, R_i)$ where R_i is the right censoring times and the censoring indicator $\Delta_i = 1(T_i \geq R_i)$. To deal with the censoring, we follow [Portnoy \[6\]](#) that proposed reweighting of the censored observations for the Kaplan-Meier estimator. He proposed estimation methods for this model that accounted for the almost inevitable presence of censoring and built it upon an analogy with the well-known univariate Kaplan-Meier estimator. Portnoy expressed the quantiles of

Kaplan-Meier distribution function $\hat{F}(t) = 1 - \hat{S}(t)$ as solutions to a weighted optimization problem in which the weight $w_i(\tau)$ associated with censored observations are split into two parts. A part of the weight associated with

$$w_i(\tau) = \frac{\tau - \tau_i}{1 - \tau_i}$$

each censored observation is left in its initial position of the censoring time with weight and the other part is shifted to the right towards positive infinity ($+\infty$) with weights $(1 - w_i(\tau))$ [Koenker \[3\]](#). By choosing different values of τ between 0 and 1, we can obtain distinct sets of estimated quantile regression coefficients and also predict the conditional quantiles of outcomes based on the given covariates. This method has been implemented using the quantreg package in R software.

7. THE LOG-RANK TEST

The log-rank test is used to compare two or more survival curves of the null hypothesis of a common survival curve; $H_0: S_1(t) = S_0(t)$

Given two groups, the log-rank statistic is based on the summed observed minus expected score for a given group and its variance estimate. Taking G to be the number of groups being compared, the test statistic is given in [Equation 13](#) as:

$$Z = \frac{\sum_{j=1}^E (O_j - E_j)}{\sqrt{\sum_{j=1}^E V_j}} \sim N(0,1) \quad \text{under } H_0 \tag{13}$$

Where O_j ; is the observed number of events is, E_j is the expected number of events and V_j is the variance of the observed number of events.

8. METHODOLOGY

In this analysis the ‘event’ considered here is ‘discharge’ from the hospital at which time covid-19 is assumed resolved. The data set for this work relates to 641 patients diagnosed with covid-19 in Rivers State Nigeria from May 2020 to August 2020. In this covid-19 research, we estimate and infer the relationship between outcomes, say, time a covid-19 patient is admitted into the hospital to the time when he/she is discharged from the hospital (denoted by $T(\text{survival time})$), and various covariates, like sex, age, fever, anosmia (absence of taste and smell), comorbidity (presence of hypertension and or biabetes) and dry-cough. The individuals admitted in the hospital and their data were available up until the end of August 2020, by which time 9 (1.4%) cases were censored while 632 (98.6%) cases were resolved and discharged from the hospital. We considered the survival analysis situation with right censoring where patients are followed until censoring occurs at time R_i which follows the continuous

distribution of the random variable T. Each patient has a censoring time R_i (time to lose to follow up or death before discharged). The sample of right-censored survival data of size $n = 641$ consists of three factors;

$$\{T_i^*, \Delta_i, X_i\}, i = 1, \dots, n$$

Where $\Delta_i(\text{censoring indicator}) = I(T_i \geq R_i), T_i^* \min(T_i, R_i)$ and X_i are the covariates.

Algorithm:- Kaplan-Meier quantiles can be computed by solving, Equation 12 starting at $\tau = 0$, Until we encounter a censored observation. Once a censored observation is encountered, i.e. when $\hat{\beta}(\tau) = T_i$, with $\Delta_i = 0$, we split T_i into two parts:

$$R_i^{(1)} = T_i \quad \text{with weight } w_i(\tau) = \frac{\tau - T_i}{1 - \tau_i} \quad (\text{the censored part left at its initial position}).$$

$$R_i^{(2)} = T_\infty = \infty \quad \text{with weight } 1 - w_i(\tau) \quad (\text{the censored part shifted to the right towards infinity}).$$

Then representing the mass set of censored observations encountered up to τ by $K(\tau)$ we can solve the Portnoy’s regression quantile estimator $\hat{\beta}(\tau)$ given in Equation 14 as;

$$\hat{\beta}(\tau) = \text{argmin}_\beta \sum_{i \notin k(\tau)} (\rho_\tau(T_i - X_i' \beta)) + \sum_{i \in k(\tau)} (w_i(\tau) \rho_\tau(T_i - X_i' \beta) + (1 - w_i(\tau)) \rho_\tau(T_\infty - X_i' \beta)) \tag{14}$$

With our model given in Equation 15 as;

$$Q_{\tau|X}(\tau|X) = \beta_0(\tau) + \beta_1(\tau) \text{sex} + \beta_2(\tau) \text{age} + \beta_3(\tau) \text{fever} + \beta_4(\tau) \text{anosmia} + \beta_5(\tau) \text{drycough} + \beta_5(\tau) \text{comorbidity} \tag{15}$$

The log-rank test was applied to compare survival curves for the different covariates considered in this work.

9. RESULTS

From the life table.1, it is seen that the marginal probability of covid-19 cases being resolved is 0.299 and the number of resolved covid-19 cases peaked between 6-7 days. The hazard rate also confirms this by telling us that the covid-19 resolved rate was highest from the 7th to the 8th day. The probability of survival is approximately 1 from the first day of admission to the next 3 days, and drops as the number of days increases.

The time interval is given in days. The graph shows an increasing hazard rate (hazard rate is the theoretical measure of the probability of occurrence of an event per unit time at risk, in this paper the event is covid-19 being resolved) which implies that as time of hospitalization increases and the symptoms drops (basically due to treatment) and the patient’s potential of covid-19 being resolved increases.

Table-1. Life Table.

Time interval	N	events	dropouts	At risk	Hazard Rate	Survival rate	Cumulative Hazard rate	Marginal Probability of event
[0,1)	353	3	1	352.5	0.00851	0.99148	0.00851	0.00851
[1,2)	349	3	3	347.5	0.00863	0.98293	0.01714	0.00856
[2,3)	343	7	1	342.5	0.02043	0.96284	0.03758	0.02009
[3,4)	335	44	1	334.5	0.13153	0.83619	0.16913	0.12665
[4,5)	290	54	0	290.0	0.18621	0.68048	0.35533	0.15570
[5,6)	236	70	0	236.0	0.29661	0.47865	0.65193	0.20184
[6,7)	166	104	0	166.0	0.62651	0.17877	1.27844	0.29987
[7,8)	62	61	1	61.5	0.99187	0.00145	2.27031	0.17732
[8,9)	288	71	0	288.0	0.2465	0.3429	0.9683	0.1122
[9,10)	217	52	0	217.0	0.2396	0.2607	1.2080	0.0822
[10,11)	165	65	1	164.5	0.3951	0.1577	1.6031	0.1030
[11,12)	99	44	0	99.0	0.4444	0.0876	2.0476	0.0701
[12,13)	55	11	0	55.0	0.2000	0.0701	2.2476	0.0175
[13,14)	44	17	0	44.0	0.3864	0.0430	2.6339	0.0271
[14,15)	27	8	1	26.5	0.3019	0.0300	2.9358	0.0130

Table-2. Survival Quantile Regression Model Results.

$\tau = 0.25$					
	Coefficients	Exp(coeff.)	Std error	T-value	P(> t)
Intercept	12.0000	16.2754e+4	84.4991	14.119	0.0000*
Sex	0.0000	1.0000	0.0000	0.6231	0.5333
Fever	2.0000	7.3891	0.0674	0.2968	0.0000*
Anosmia	-4.0000	0.0183	0.8508	-4.7013	0.0000*
Cough	0.0000	1.0000	0.0000	-1.9478	0.8456
Age	0.0000	1.0000	0.0000	2.2085	0.0027*
Comorbidity	-2.0029	0.1349	0.8441	-2.3727	0.0177*
$\tau = 0.50$					
	Coefficients	Exp(coeff.)	Std error	T-value	P(> t)
Intercept	13.4056	66.3709e+4	2.1096	6.3546	0.0000*
Sex	0.0000	1.0000	0.0000	-0.08770	0.9301
Fever	3.0000	20.0855	0.5630	5.3286	0.0000*
Anosmia	-3.2650	0.0382	0.3648	-8.9497	0.0000*
Cough	-0.8594	0.4234	1.1395	-0.7542	0.4507
Age	0.0000	1.0000	0.0000	0.4697	0.6386
Comorbidity	-2.1406	0.1176	2.2658	-0.9447	0.3448
$\tau = 0.75$					
	Coefficients	Exp(coeff.)	Std error	T-value	P(> t)
Intercept	20.488	79.0360e+6	0.7232	28.3297	0.0000*
Sex	-0.4278	0.6519	0.2994	-1.4290	0.1530
Fever	2.6107	13.6086	0.2746	9.5062	0.0000*
Anosmia	-9.9084	0.0001	0.4713	-21.0225	0.0000*
Cough	0.3437	1.4101	0.2881	1.1929	0.2329
Age	0.0229	1.0231	0.0167	1.3695	0.1709
Comorbidity	-1.6177	0.1983	0.4591	-3.5234	0.0004*
$\tau = 0.95$					
	Coefficients	Exp(coeff.)	Std error	T-value	P(> t)
Intercept	17.5008	39.8566e+6	1.1593	15.0960	0.0000*
Sex	0.0580	1.0597	0.5207	0.1115	0.9113
Fever	2.7852	16.2030	0.9426	2.9547	0.0031*
Anosmia	-3.2078	0.0404	2.7159	-1.1811	0.2376
Cough	-0.3498	0.7048	0.9171	-0.3814	0.7029
Age	0.0446	1.0456	0.0437	1.0191	0.3082
Comorbidity	0.5989	1.8201	0.9249	0.6475	0.5173

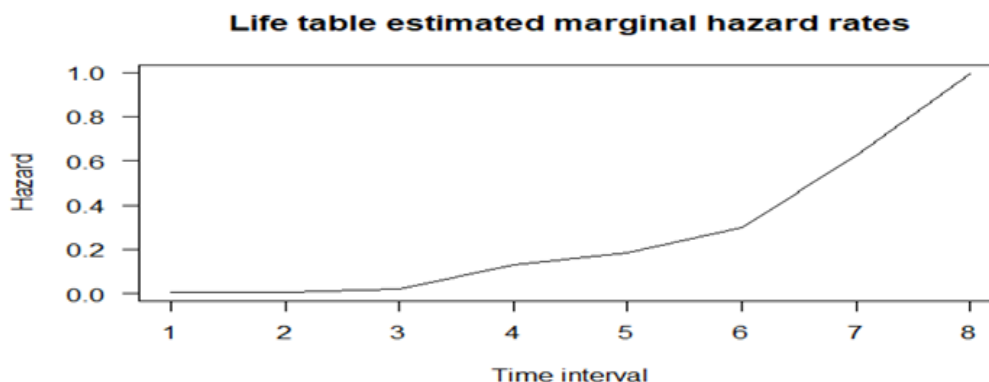


Figure-1. Margiinal hazard curve.

Table 2 presents the results of the survival quantile regression model in Equation 14, the results show that at the 25th quantile all the covariates were significant except sex and cough which tells us that on admission into the hospital the survival time of the patients depended on their age and the presence of anosmia, comorbidity and fever. But by the mid survival time only anosmia and fever were seen to be significant and by the 95th quantile i.e as the length of stay in the hospital increases the effect of anosmia disappears while the fever effect was still seen to be significant. The exponentiated coefficients known as the hazard ratio shows the size of the effect of the covariates. The result shows that the size of the effect of fever increases the hazard by 7.9389 at the 25th quantile and peaks at the 50th quantile by 20.0855 and drops at the 75th quantile to 13.6086 but increases a bit at the 95th quantile to 16.2030, which means that having fever is associated with longer stay in the hospital. We also noticed that the effect of anosmia and comorbidity at the 75th quantile were significant but the sizes of the effects were minimal though comorbidity has a bigger effect than anosmia.

10. RESULTS BASED ON SEX

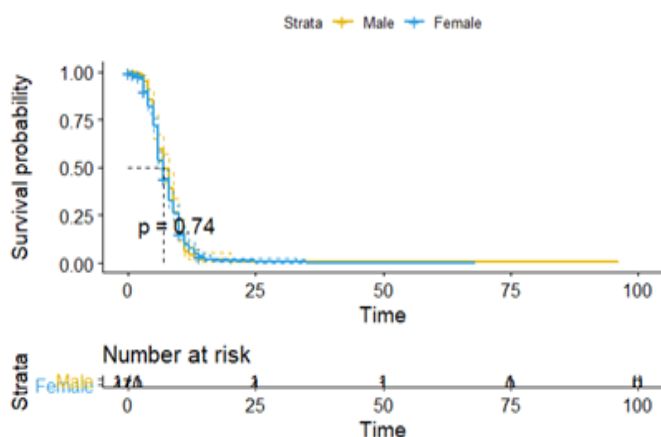


Figure-2a. Kaplan-Meier Survival Curve.

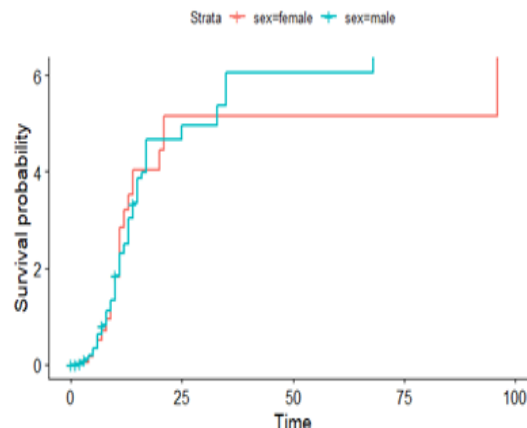


Figure-2b. Cumulative Hazard Plot.
Source: survival and survminer packages in R software

Based on the graph both male and female patients have the same probability of survival and the probability of survival is 0.74. The cumulative hazard plot also confirms that the survival time of both sex are similar.

Table-3. The Log-rank test results:

Sex	N	Observed	Expected	$(O_j - E_j)^2/E$	$(O_j - E_j)^2/V_j$
Female	171	171	174	0.0628	0.112
Male	470	461	458	0.0239	0.112

Note: Chisq= 0.1 on 1 degrees of freedom, p= 0.7.

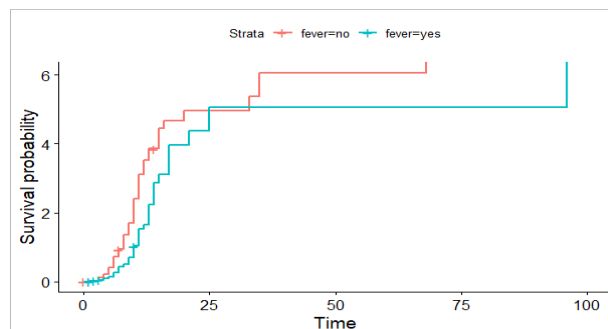
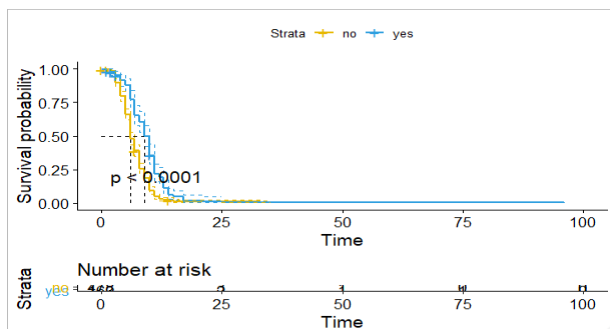
The test result is not significant; therefore the probability of survival of male and female are not different, agreeing with the graph.

Table-4. The Confidence interval of survival time:

Sex	N	Events	Median	0.95 LCL	0.95 UCL
Female	171	171	7	7	8
Male	470	461	7	6	7

The table shows that female patients' survival time was between 7-8 days with a median of 7 days while men patients' survival time was between 6-7 days also with a median of 7 days which is the same for both.

11. RESULTS BASED ON FEVER



Source: survival and survminer packages in R software

The plot shows that the probability of survival of patients with fever and those without fever is different. The patients without fever survive at a shorter time interval than the patients with fever with a probability of 0.0001 more than the patients with fever. The cumulative hazard plot also shows the gap in survival between those with fever and those without fever.

Table-5. The Log-rank test results.

Fever	N	Observed	Expected	$(O_j - E_j)^2/E$	$(O_j - E_j)^2/V_j$
No fever	478	473	393	16.4	59.7
Yes fever	163	159	239	26.9	59.7

Note: Chisq= 59.7 on 1 degrees of freedom, p= 1e-14

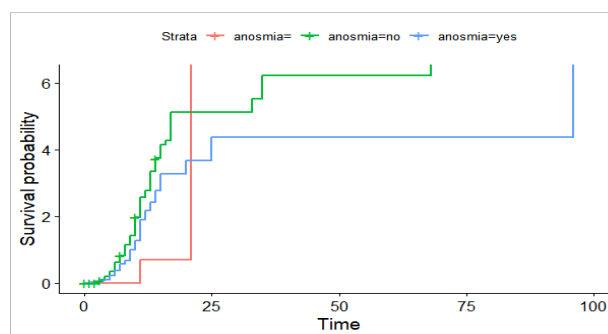
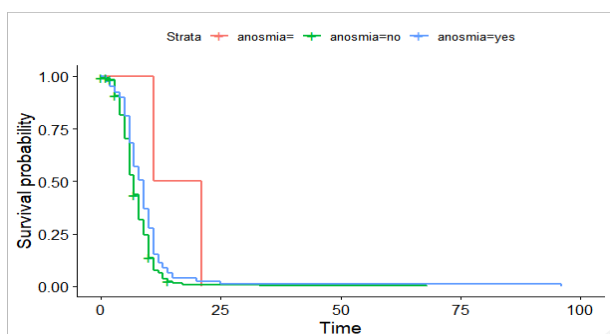
This test confirms what the graphs in Figure 4 & 5 shows, that there is a significant difference between the probability of survival for patients with fever and those without fever.

Table-6. The Confidence interval of survival time.

Fever	N	events	median	0.95 LCL	0.95 UCL
No fever	478	473	6	6	7
Yes fever	163	159	9	9	10

The patients that had fever were hospitalized between 9-10 days with a median of 9 days while those without fever were hospitalized between 6-7 days with a median of 6 days before the covid-19 issue was resolved. This shows that those with fever stayed longer on admission than those without fever.

12. SURVIVAL ANALYSIS BASED ON ANOSMIA



The survival curve in Figure (6a) shows that there is a significant difference in the survival times between the patients that have anosmia and those that do not have anosmia. The cumulative curve hazard in Figure (6b) above also confirms the difference.

Table-7. The Log-rank test results.

Anosmia	N	Observed	Expected	$(O_j - E_j)^2/E$	$(O_j - E_j)^2/V_j$
Yes anosmia	79	79	103.36	5.74	9.13
No anosmia	560	551	522.58	1.55	11.97
Non response	2	2	16	11	NA

Note: Chisq= 13.6 on 2 degrees of freedom, p=0.001.

The log-rank test confirms that there is a significant difference in survival time between those that have anosmia and those that do not have anosmia.

Table-8. The Confidence interval of survival time.

Anosmia	n	Events	Median	LCL	UCL
Yes anosmia	79	79	9	7	9
No anosmia	560	551	7	6	7
Non response	2	2	16	NA	NA

The table shows that patients who had anosmia were hospitalized between 7-9 days with a median of 9 days while those without anosmia were hospitalized between 6-7 days with a median of 6 days before the Covid-19 issue was resolved. This shows that those with anosmia stayed longer on admission than those without anosmia.

13. SURVIVAL BASED ON COUGH

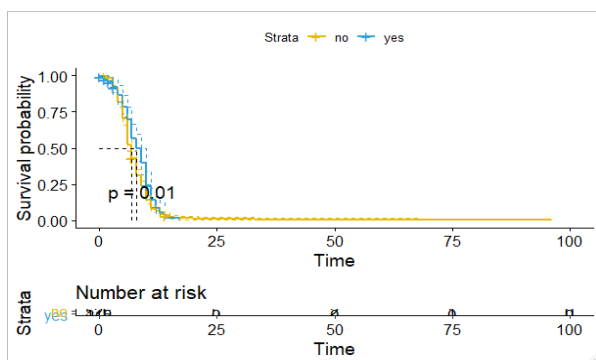


Figure-5a. Kaplan-Meier Survival Curve.

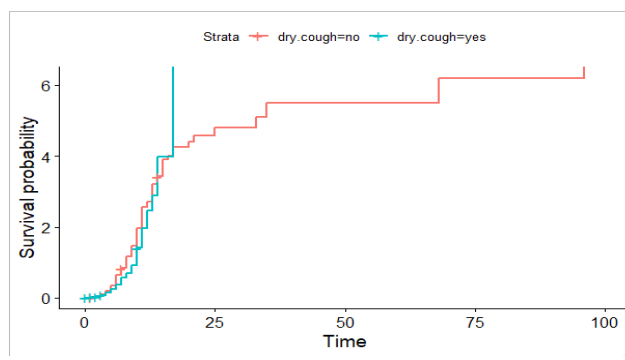


Figure-5b. Cumulative Hazard Plot.

The plot in Figure (5a) shows that the probability of survival of patients with cough and those without cough is different. The patients without cough survive at a shorter time interval than the patients with cough with a probability of 0.01 more than the patients with cough. The cumulative hazard plot of Figure 5b also confirms that.

Table-9. The Log-rank test results.

Cough	N	Observed	Expected	$(O_j - E_j)^2/E$	$(O_j - E_j)^2/V_j$
No cough	526	522	499	1.06	6.61
Yes cough	115	110	133	3.97	6.61

Note: Chisq= 6.6 on 1 degrees of freedom, p= 0.01.

The log-rank test confirms that there is a significant difference in survival time between those that have cough and those that do not have cough.

Table-10. The Confidence interval of survival time.

Cough	N	Events	median	LCL	UCL
No cough	526	522	7	6	7
Yes cough	115	110	8	7	9

The patients that had cough were hospitalized between 7-9 days with a median of 8 days while those without cough were hospitalized between 6-7 days with a median of 7 days before the covid-19 issue was resolved. This shows that those with cough stayed longer on admission than those without cough.

14. SURVIVAL BASED ON COMORBIDITY

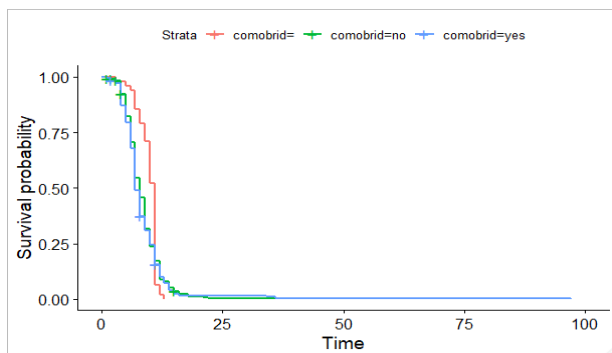


Figure-6a. Kaplan-Meier Survival Curve.

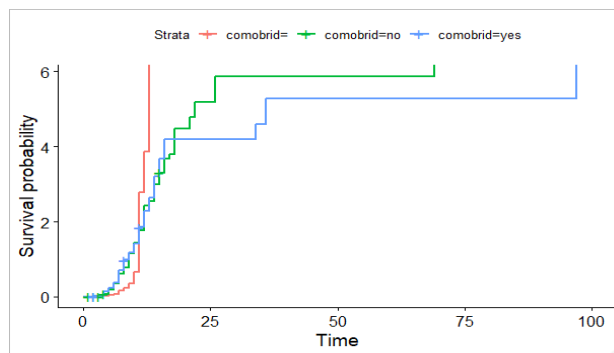


Figure-6b. Cumulative Hazard Plot.

The plot in Figure 6a shows that the probability of survival of patients with hypertension and or diabetes and those without either of them is not different. The cumulative hazard plot of Figure 6b also confirms this claim.

Table-11. The Log-rank test results.

Comorbidity	N	Observed	Expected	$(O_j - E_j)^2 / E$	$(O_j - E_j)^2 / V_j$
No comorbidity	385	380	377.6	0.0148	0.0474
Yes comorbidity	208	208	193.2	0.6071	1.1310
No response	48	48	61.2	2.8432	4.1358

Chisq= 4.5 on 2 degrees of freedom, p= 0.1.

The log-rank test confirms that there is no significant difference in survival time between those that have hypertension and or diabetes and those that do not have either or both.

Table-12. The Confidence interval of survival time.

Comorbidity	N	Events	median	LCL	UCL
No comorbidity	385	380	8	7	9
Yes comorbidity	208	204	7	7	8
No response	48	48	11	11	11

The patients that had hypertension and or diabetes were hospitalized between 7-8 days with a median of 7 days while those without hypertension and or diabetes were hospitalized between 7-9 days with a median of 7 days before the covid-19 issue was resolved. This confirms the results of the log rank test that those with or without the presence of comorbidity had the same survival time.

15. CONCLUSION

In conclusion we state that;

- Most covid-19 cases were resolved after about 7-8 days of being admitted.
- As survival time increases the anosmia symptoms and comorbidity effect disappears and the probability of being discharged (case resolved) increases.
- Sex and cough have no significant effect on survival time of covid-19 patients.
- Fever, comorbidity, age and anosmia have significant effects on the survival time of covid-19 patients at the 25th quantile.
- Fever had a significant effect at all quantiles, with effect size of 7.3890 at the 25th quantile, effect size of 20.0855 at the 50th quantile, effect size of 13.6086 at the 75th and effect size of 16.2030 at 95th quantile.
- Male and female patients have the same probability of survival and their probability of survival is 0.74
- Patients with comorbidity and those without have the same probability of survival, and their survival time is between 7-9 days.
- The patients without fever survive at a shorter time interval than the patients with fever with a probability of 0.0001 more than the patients with fever.
- The survival time of patients with fever is between 9-10 days while those without fever are between 6-7 days. This shows that those with fever stayed longer on admission than those without fever.

- There is a significant difference in the survival times between the patients that have anosmia and those that do not have anosmia.
- Patients that had anosmia were hospitalized between 7-9 days while those without anosmia were hospitalized between 6-7 days before the covid-19 issue was resolved. This shows that those with anosmia stayed longer on admission than those without anosmia.
- The patients without cough survive at a shorter time interval than the patients with cough with a probability of 0.01 more than the patients with cough.
- Patients with cough were hospitalized between 7-9 days while those without cough were hospitalized between 6-7 days before the covid-19 issue was resolved. This shows that those with cough stayed longer on admission than those without cough.
- At the onset of covid-19, the survival time of patients depends on fever, age, comorbidity and anosmia. But as patients are being treated the effects of these factors becomes non-significant except the effect of fever.

Funding: This study received no specific financial support.

Competing Interests: The authors declare that they have no competing interests.

Acknowledgement: Both authors contributed equally to the conception and design of the study.

REFERENCES

- [1] D. R. Cox, "Regression models and life-tables," *Journal of the Royal Statistical Society: Series B (Methodological)*, vol. 34, pp. 187-202, 1972.
- [2] J. Buckley and I. James, "Linear regression with censored data," *Biometrika*, vol. 66, pp. 429-436, 1979.
- [3] R. Koenker, *Quantile regression. Econometric society monographs* vol. 38. London: Cambridge U. Press, 2005.
- [4] J. L. Powell, "Censored regression quantiles," *Journal of Econometrics*, vol. 32, pp. 143-155, 1986.
- [5] Z. Ying, S.-H. Jung, and L.-J. Wei, "Survival analysis with median regression models," *Journal of the American Statistical Association*, vol. 90, pp. 178-184, 1995.
- [6] S. Portnoy, "Censored regression quantiles," *Journal of the American Statistical Association*, vol. 98, pp. 1001-1012, 2003.
- [7] L. Peng and Y. Huang, "Survival analysis with quantile regression models," *Journal of the American Statistical Association*, vol. 103, pp. 637-649, 2008.
- [8] R. Koenker and O. Geling, "Reappraising medfly longevity: A quantile regression survival analysis," *Journal of the American Statistical Association*, vol. 96, pp. 458-468, 2001.

Views and opinions expressed in this article are the views and opinions of the author(s), International Journal of Mathematical Research shall not be responsible or answerable for any loss, damage or liability etc. caused in relation to/arising out of the use of the content.