Results of Statistical Analyses

Predictors of mortality of hospitalized patients with laboratory-confirmed severe acute respiratory syndrome coronavirus 2 in Nigeria: A retrospective analysis of high burden states

October 12, 2021

Contents

List of Tables

List of Figures

Background

The current COVID 19 pandemic has resulted in a high number of deaths and associated disruption of public health and socioeconomic activities of countries and populations. As at 9th of April 2021, Nigeria had recorded 163,581 confirmed cases, with 150,005 cases discharged, 7,518 cases on admission and 2,058 deaths with a significant increase in the number of confirmed cases and concurrent deaths since the beginning of the second wave of infections. Most of these deaths have been in-hospital. Little is known about the deaths in the community due largely to poor community records of deaths and other vital statistics. The weak Nigerian health system is thought to have led to a considerably high in-hospital mortality especially in settings where critical care services and resources are scarce. Underlying morbidities such as malnutrition, anemia, HIV/AIDs, and chronic respiratory conditions, diabetes and heart failure have been shown to be important contributors to high global mortality in the current COVID 19 pandemic.

Even though several global reports have been written about the impact of the COVID 19 pandemic on clinical outcomes especially the attendant morbidity and mortality, little is known about the situation in Nigeria. Understanding the relative contributions and probable mechanism through which sociodemographic, clinical and laboratory factors relate to the high in-hospital mortality in Nigeria could help us identify weak points within the health system that can be improved upon for current and future responses. In addition, such review will also help us plan and prioritize health infrastructure and resource allocation for better health outcomes in our public health system.

The aim of this study was to describe clinical characteristics and factors associated with mortality and time to mortality for patients hospitalized with laboratory-confirmed severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) infection in Nigeria.

Data Analysis Plan

Study Setting

Nigeria is Africa's most populous country. We studied patients with laboratory-confirmed SARS-CoV-2 infection from 28 COVID 19 treatment centers across 15 states in Nigeria including the Federal Capital Territory (see Figure 1).

Study Design

This study was a retrospective cross-sectional study.

Data Collection

We extracted data from the World Health Organization (WHO) nCOV database platform (OpenClinica Electronic Data Capture [EDC] System) for 28 sites across 15 states in Nigeria (see Figure 1). We retrieved data on date of admission, enrollment date, demographic data (patients' sex, age, health worker), data on pre-existing morbidity (chronic cardiac disease, hypertension, chronic pulmonary disease, asthma, chronic kidney disease, chronic neurological disorder, HIV, diabetes, current smoking, tuberculosis, asplenia, malignant neoplasm), pre-admission and chronic medication (ACE inhibitors, angiotensin II receptor

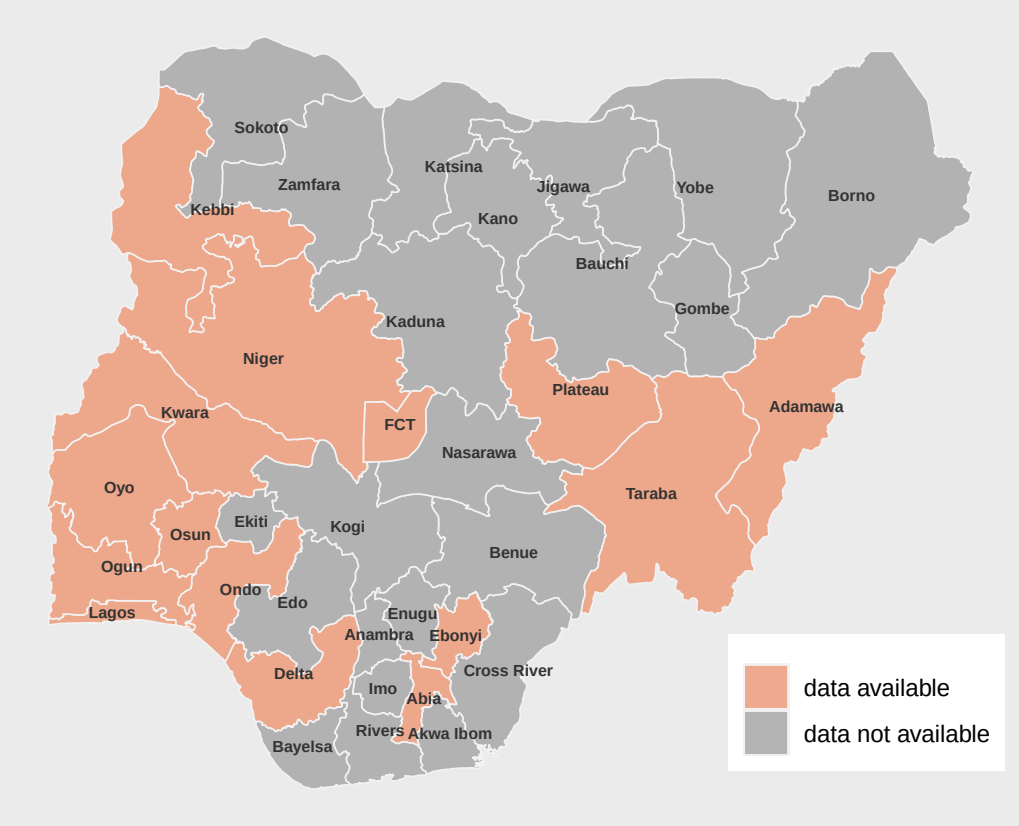

Figure 1: States with available data

blockers, non-steroidal anti-inflammatory drugs in the 14 days prior to admission), medications received on day of admission or following (antiviral, corticosteroid, chloroquine/hydroxychloroquine), admitted to ICU or HDU, supplemental oxygen given, use of non-invasive ventilation, use of invasive ventilation, outcome (discharged alive, hospitalized, transfer to other facility, death, palliative discharge), and outcome date. Following data cleaning, data recoding and removal of missing observations we were left with 3,462 records for data analysis. Two states (Delta and Taraba) were excluded because the states each had less than 10 records: Delta (6), Taraba (1).

Study Variables

The primary study outcome was death for a PCR confirmed infection with SARS-COV-2 virus. The secondary study outcome was time from enrollment to death. We coded the primary study outcome as 1 if the patient died or 0 if outcome was censored. Time following admission was right-censored at 84 days post-enrollment. The main explanatory variable was severity of disease. We defined severity of disease as a composite score derived from any or all of:

1. Received supplemental oxygen therapy on any day during hospitalization,
2. Admitted to the ICU or HDU at any point during hospital stay,
3. Received non-invasive ventilation (CPAP/BIPAP) on any day during hospitalization, and
4. Received invasive ventilation (mechanical ventilator) on any day during hospitalization

These variables were assigned a score of 1 if reported in the database or 0, otherwise. The severity score was calculated as the sum of these individual scores with a range of 0 to 4. The other explanatory variables were demographic (age, sex, healthcare worker), pre-existing morbidity (), malaria, HIV infection, pre-admission and chronic medications, and medications received on admission or following admission. These explanatory variables were coded as 1 if reported or 0 if not reported or missing.

Statistical Analysis

Power calculation

The relationship between the log odds of the mortality and *k* explanatory variables may be modeled thus:

$$\log\left(\frac{p}{1-p}\right) = \beta_1 x_1 + \beta_2 x_2 + \cdots + \beta_k x_k$$

The main explanatory covariate was severity score assumed to be normally distributed based on the central limit theorem. Given two-sided testing of $H_0 : \beta_i = 0$ on the log scale versus $H_1 : \beta_i \neq 0$ the minimum sample size is then, according to Hsieh et al.,[1] given by:

$$n = \frac{(z_{1-\alpha/2} + z_\gamma)^2}{p(1-p)(1-\rho_i^2)\hat{\beta}^2}$$

with level α and power γ, the standard deviation of the predictor x_i is σ_{x_i}, p the marginal prevalence of the outcome, ρ_i^2 is the multiple correlation of x_i with all the other predictors (i.e. the R^2), $z_{1-\alpha/2}$ and z_γ are quantiles of the standard normal distribution corresponding to level and power, $\hat{\beta}^2$ is the effect size (i.e. the log odds ratio):

$$\hat{\beta} = \log\left(\frac{p_1/1-p_1}{p_2/1-p_2}\right)$$

and p_1, p_2 are the event rates at the mean of the severity score, and one standard deviation above the mean severity score respectively.

Therefore the power is calculated as

$$\gamma = 1 - \Phi\left(z_{1-\alpha/2} - |\hat{\beta}|\sqrt{np(1-p)(1-\rho_i^2)}\right)$$

Where Φ is the standard normal cumulative distribution function.

Using G*Power[2] with α = 0.05, $\hat{\beta}$ = 1.5, p = 0.05, R^2 = 0.25 and assumed mean severity score of 2 with standard deviation of 1, we determined that the number of subjects of 3,462 had greater than 95% power for a two-tailed hypothesis test (see Figure 2).

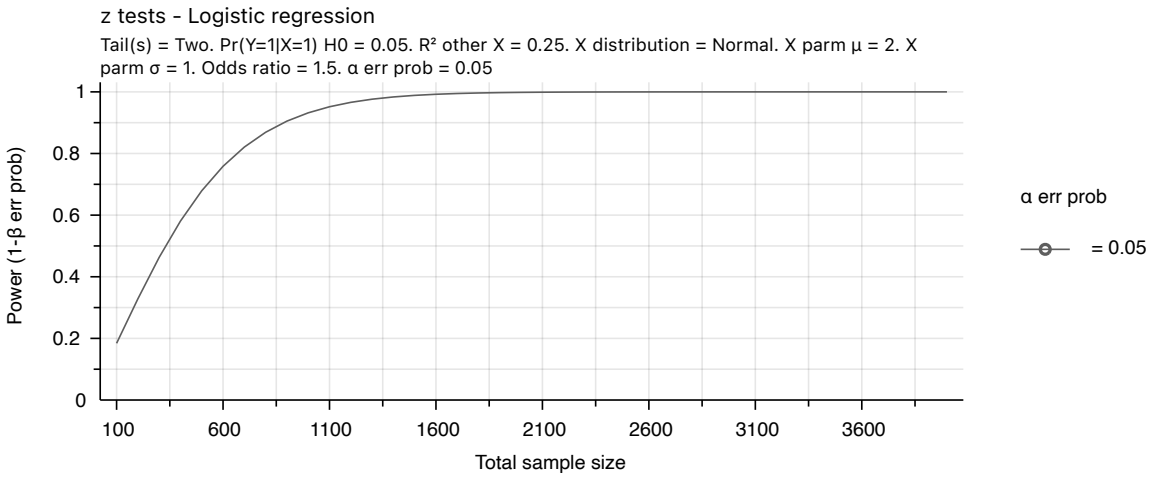

Figure 2: Plot of sample sizes versus power

Regression modeling

The term multilevel refers to individuals at a lower level who are nested within spatial units at higher levels.[3] Multilevel methods are suitable for the statistical analysis of data with a nested structure. Multilevel regression modeling was used for the data analysis in this study. These data were from several COVID-19 treatment sites across Nigeria (Figure 1) with clustering of patients by site. The sites in turn may be considered to be nested by facility type e.g. tertiary hospital, non-tertiary hospital, and other treatment site.

Multilevel statistical modeling incorporates models at each level of analysis into a full multilevel model. In general terms, for example, in a two-level nested model, the model at the first level is expressed as

$$y_{ij} = \beta_{0j} + \beta_1 x_{1ij} + e_{0ij}$$

Where y_{ij} is the measure of the dependent variable for the ith individual in the jth group. The term β_{0j} is a constant and is the measure of the dependent variable for the jth group, and β_1 is the fixed marginal effect of the predictor variable (x_{1ij}) on the dependent variable.

The individual or the level-1 residual term, e_{0ij}, is assumed to have a normal distribution with a mean of 0 and a variance, $\sigma^2_{e_0}$. In multilevel modeling, the coefficients at level-1 become the outcome variables at level-2. Thus, the model at level-2 can be written as

$$\beta_{0j} = \beta_0 + u_{0j}$$

meaning the mean measure of the dependent variable for the jth group is split into β_0 (the average for the dependent variable across all groups), and u_{0j}, the effect specific to the jth group, and u_{0j} can be treated similar to individual-level residuals. Combining the two equations above yields the full model. This full model is known as a random-intercepts or variance components model:

$$y_{ij} = \beta_0 + \beta_1 x_{1ij} + (u_{0j} + e_{0ij})$$

In this multilevel statistical model, the variance at level-2, $\sigma^2_{u_0}$, measures the group differences after accounting for the compositional effect of the predictor variable and thus separates the effects of the individual-level variables from the contextual differences between group-level variables.[3] Therefore using multilevel modeling we can (a) model the variation in patient-level regression coefficients across sites, (b) account for the patient- and site-level variation in estimating site-level coefficients, and (c) estimate regression coefficients for particular groups.[4]

We performed univariate analysis using the Fisher exact test, or the Wilcoxon test as appropriate to describe characteristics of patients who had non-missing outcome data (see Table 1). We also did univariate survival analysis using the Kaplan-Meier method to assess the relationship between time from enrollment to death and patient characteristics. We further fitted mixed effects logistic regression models and Cox proportional hazards regression models to assess predictors of COVID-19 hospital death. Study site and facility type were included in the models as random effects while severity score, patients' sex, age, health worker status, chronic cardiac disease, hypertension, respiratory disease (any of chronic pulmonary dis-

ease, active or previous tuberculosis, asthma), chronic kidney disease, chronic neurological disorder, HIV, diabetes, current smoking, malignant neoplasm, malaria, pre-admission and chronic medication (ACE inhibitors, angiotensin II receptor blockers, non-steroidal anti-inflammatory drugs in the 14 days prior to admission), medications received on day of admission or following admission (antiviral, corticosteroid, azithromycin, chloroquine/hydroxychloroquine) were included as fixed effects. We report the models with the smallest AIC selected by stepwise regression. All statistical analysis were conducted using R version 4.1.1.[5] Mixed effects Cox regression was done using the `coxme` package.[6]. A p value < 0.05 is considered as statistically significant.

Results

Patient characteristics

Table 1: Patient characteristics

	Died		
Variables	**No**, *N* = 3,249	**Yes**, *N* = 213	**p value**
Severity score			**<0.001**
0	2,916 (89.8%)	89 (41.8%)	
1	249 (7.7%)	71 (33.3%)	
2	78 (2.4%)	47 (22.1%)	
3	5 (0.2%)	5 (2.3%)	
4	1 (0%)	1 (0.5%)	
Days from enrollment to outcome, Median (IQR)	12 (7, 15)	2 (1, 10)	**<0.001**
Age, Median (IQR)	39 (28, 53)	61 (48, 70)	**<0.001**
Age group			**<0.001**
<18	252 (8.1%)	4 (1.9%)	
18—45	1,688 (54.4%)	39 (18.7%)	
46—65	894 (28.8%)	95 (45.5%)	
66—75	183 (5.9%)	41 (19.6%)	
>75	87 (2.8%)	30 (14.4%)	
Male	1,947 (59.9%)	152 (71.4%)	**<0.001**
Health care worker	144 (4.4%)	3 (1.4%)	**0.034**
Chronic cardiac disease (not hypertension)	19 (0.6%)	13 (6.1%)	**<0.001**
Diabetes	237 (7.3%)	62 (29.1%)	**<0.001**
Hypertension	506 (15.6%)	88 (41.3%)	**<0.001**
Current smoking	18 (0.6%)	4 (1.9%)	**0.042**
Respiratory disease	55 (1.7%)	4 (1.9%)	0.782
Chronic neurologic disease	12 (0.4%)	0 (0%)	>0.999
Malignant neopllasm	8 (0.2%)	2 (0.9%)	0.122
Chronic kidney disease	10 (0.3%)	6 (2.8%)	**<0.001**
HIV	24 (0.7%)	4 (1.9%)	0.09
Malaria	78 (2.4%)	6 (2.8%)	0.702
Antiviral	457 (14.1%)	24 (11.3%)	0.253
ACE-i/ARB	94 (2.9%)	10 (4.7%)	0.136
Azithromycin	1,515 (46.6%)	104 (48.8%)	0.534
Corticosteroid	180 (5.5%)	54 (25.4%)	**<0.001**
CQ/HCQ	968 (29.8%)	71 (33.3%)	0.275
NSAIDs	31 (1%)	3 (1.4%)	0.463
Systemic anticoagulation	58 (1.8%)	15 (7%)	**<0.001**

ACE-i/ARB = Angiotensin-converting enzyme inhibitors; NSAIDs = Non-steroidal anti-inflammatory drugs; CQ/HCQ = Chloroquine/hydroxychloroquine

Fisher's exact test; Wilcoxon rank sum test; Pearson's Chi-squared test

Table 1 shows that of the 3,462 patients with non-missing outcomes, 0 patients died representing a 0% mortality rate. Half of the deaths occurred within two days post-enrollment. Unadjusted analyses showed that increasing severity score, older age, being male, smoking, chronic cardiac disease, diabetes, hypertension, chronic kidney disease, corticosteroid use, and systemic coagulation were statistically significantly associated with increased mortality in these patients. Being a healthcare worker was statistically significantly associated decreased mortality.

Kaplan-Meier survival curves

This section includes the Kaplan-Meier survival curves of patients stratified by different variables. Similar to Table 1, the time-to-event analysis showed that increasing age, male sex, diabetes, hypertension, current smoking, chronic kidney disease, steroid use, and systemic coagulation were associated with decreased survival probability. Unlike seen in table 1, healthcare worker status, chronic cardiac disease were not statistically significantly associated with survival probability. The use of chloroquine/hydroxychloroquinne however was found to be statistically significantly associated with increased probability of survival (Figure 19).

Demographics

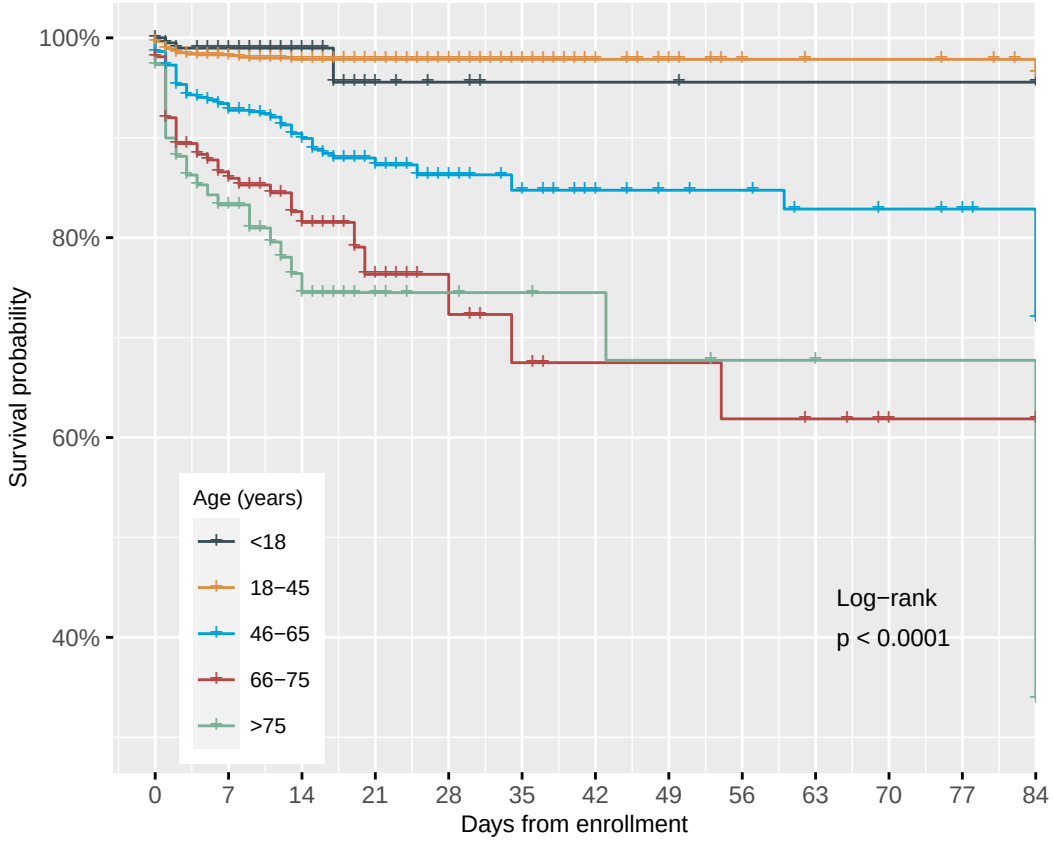

Figure 3: Survival curves of patients stratified by age

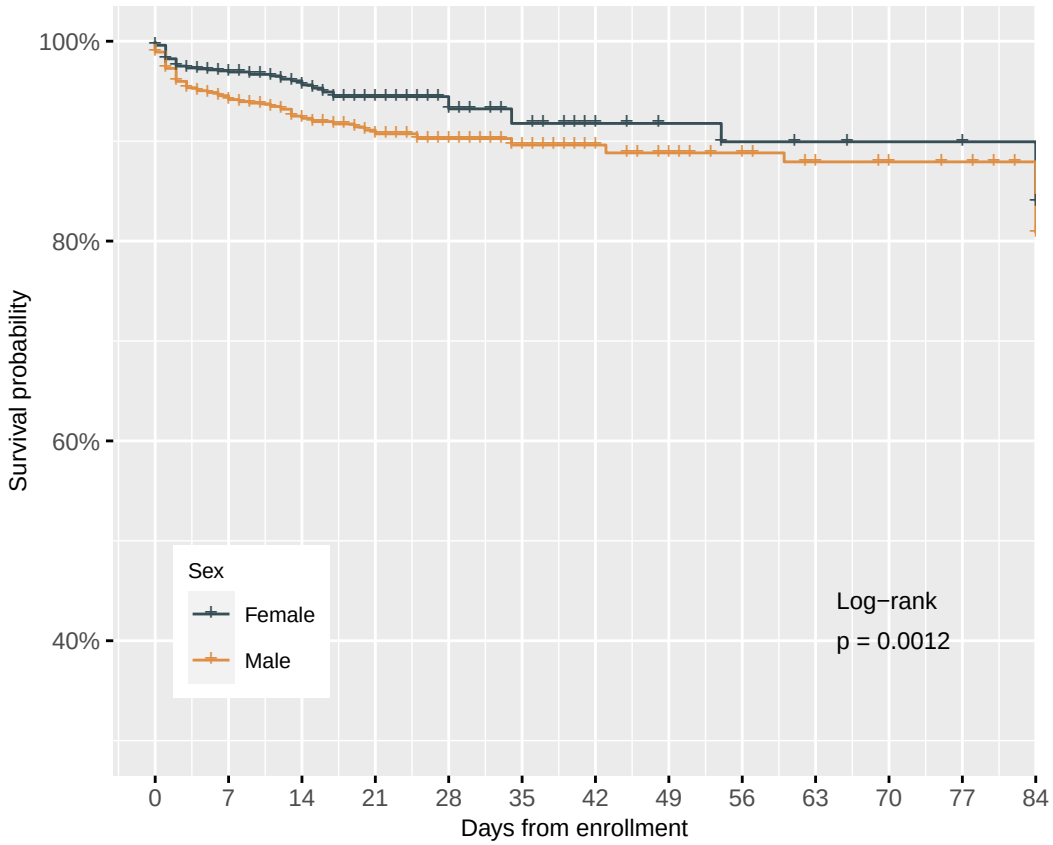

Figure 4: Survival curves of patients stratified by sex

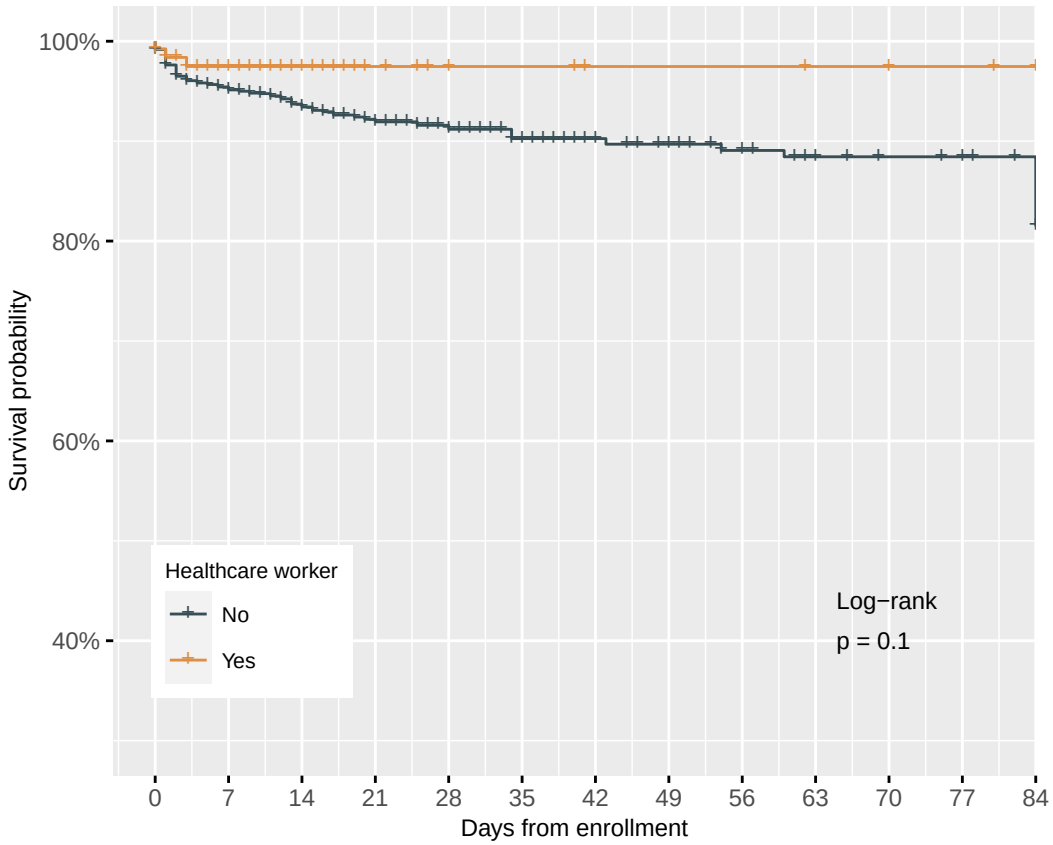

Figure 5: Survival curves of patients stratified by healthcare worker status

Comorbid conditions

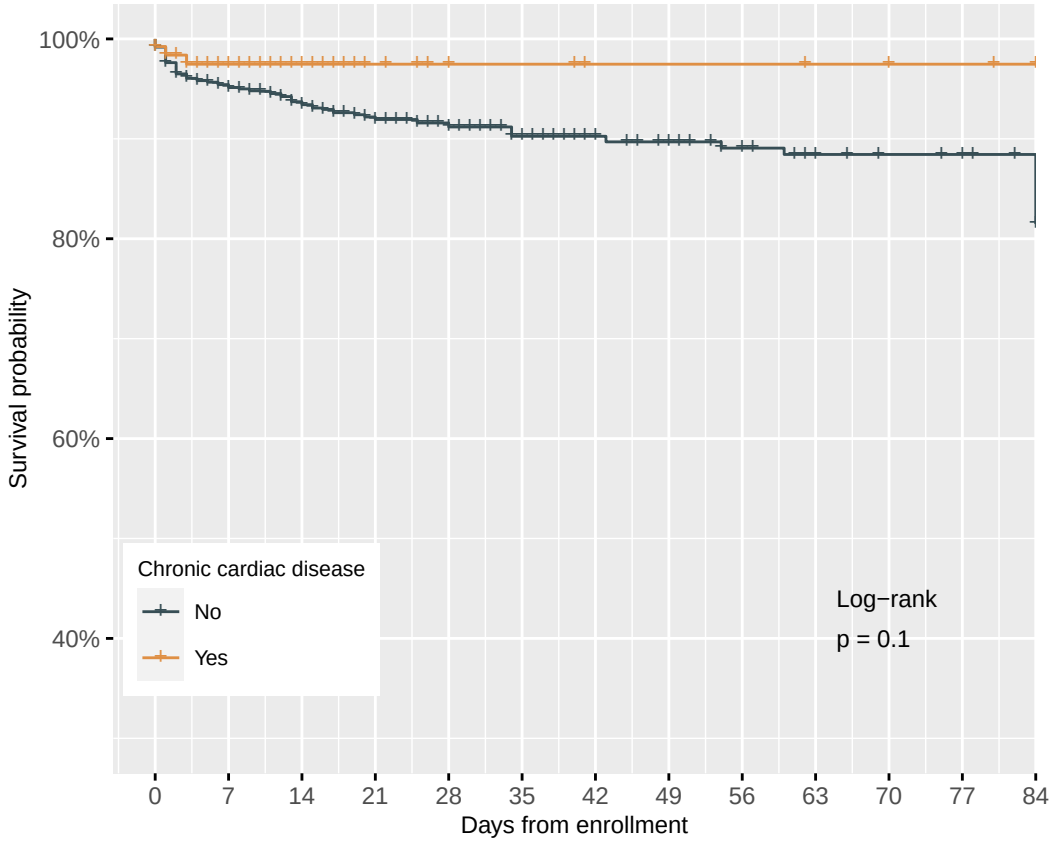

Figure 6: Survival curves of patients stratified by chronic cardiac disease status

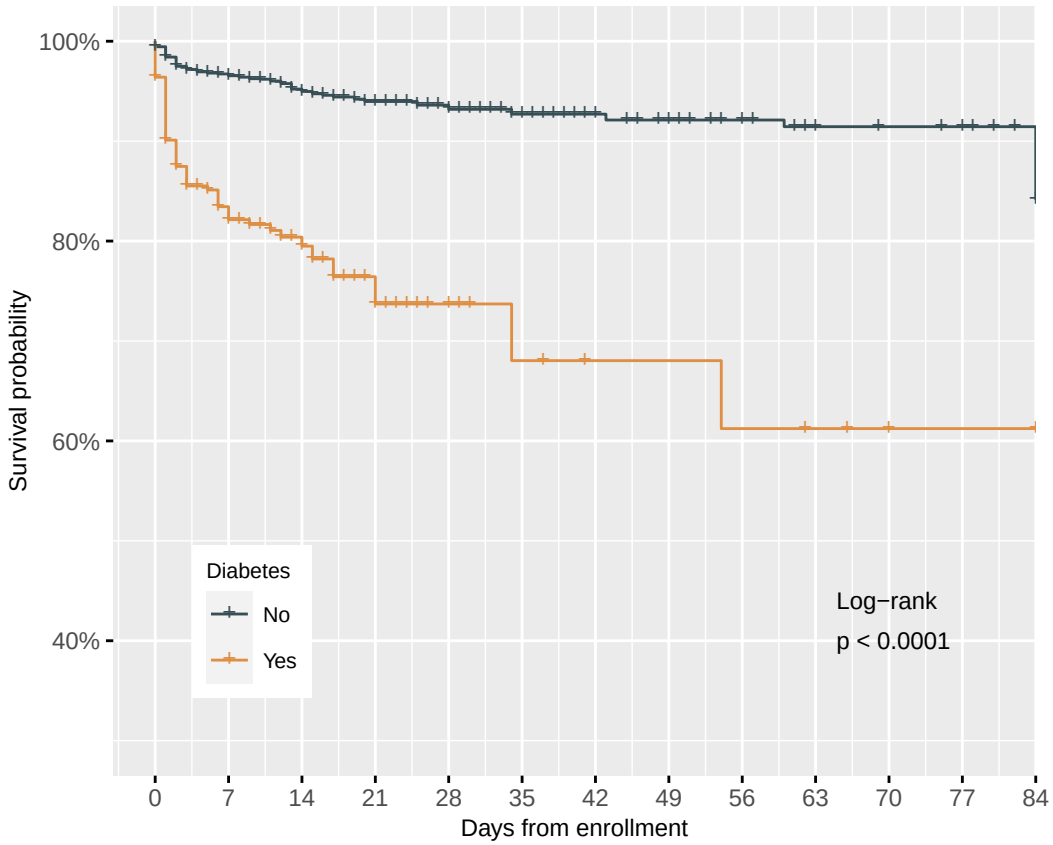

Figure 7: Survival curves of patients stratified by diabetes status

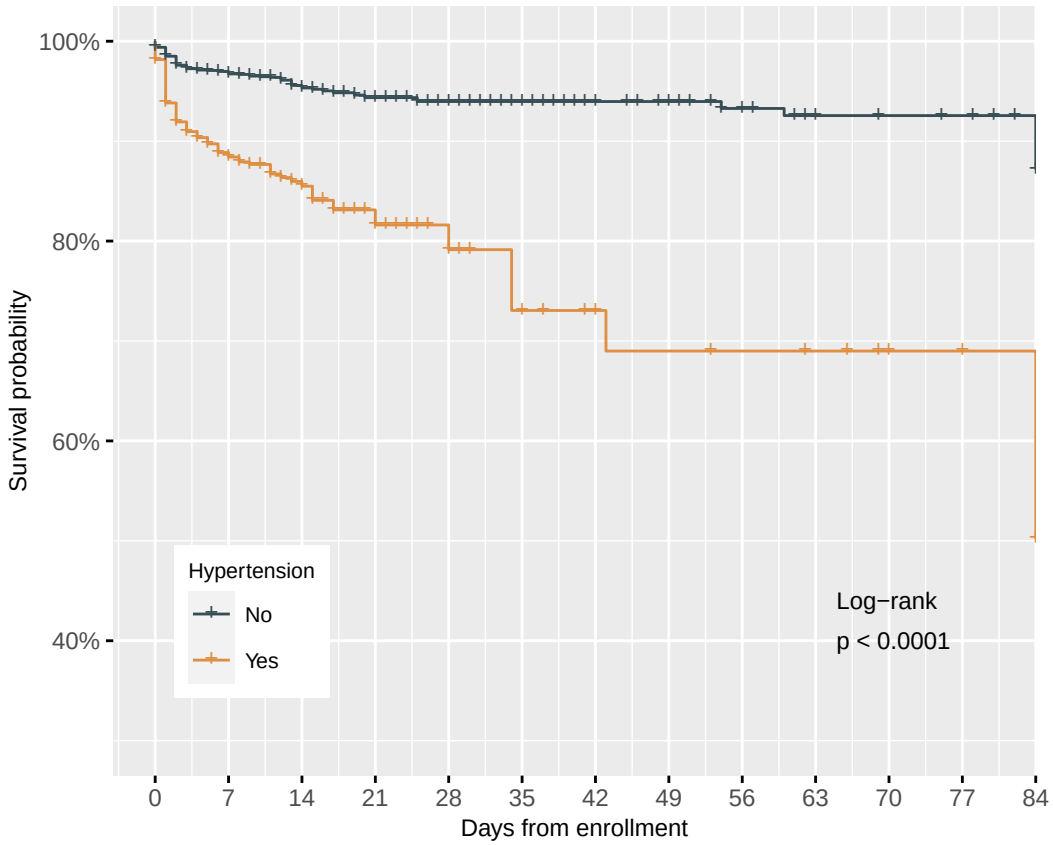

Figure 8: Survival curves of patients stratified by hypertension status

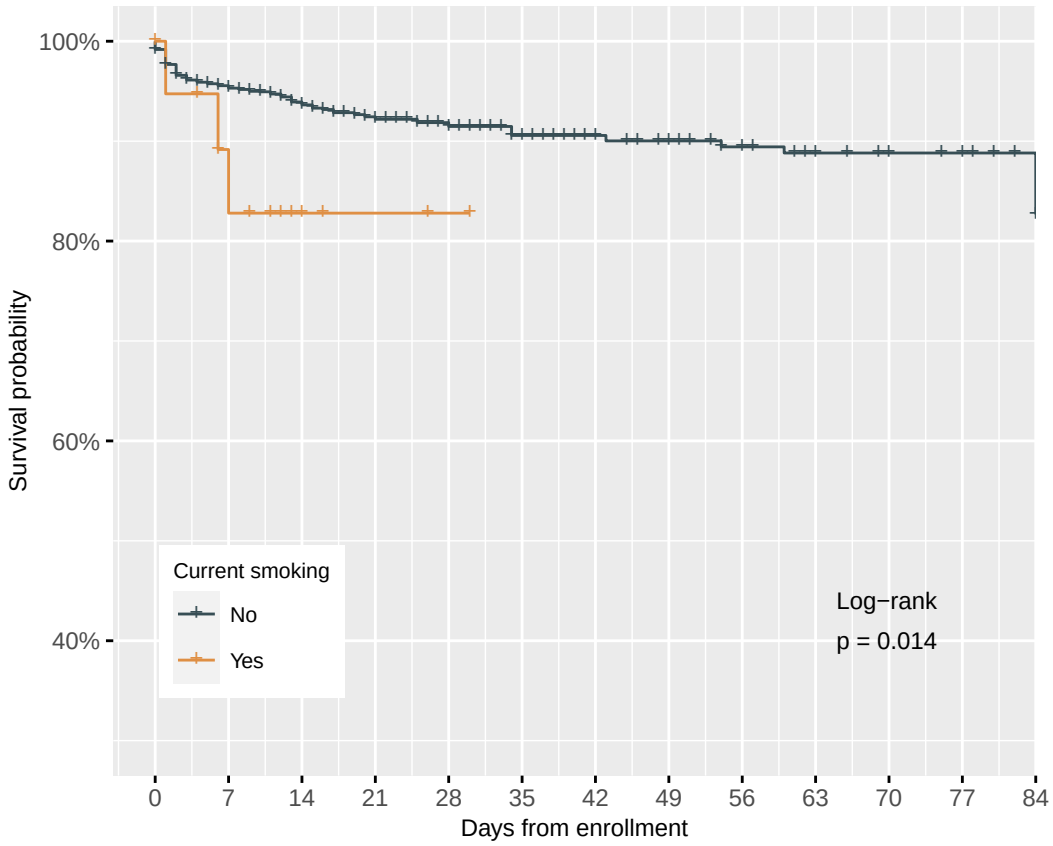

Figure 9: Survival curves of patients stratified by current smoking status

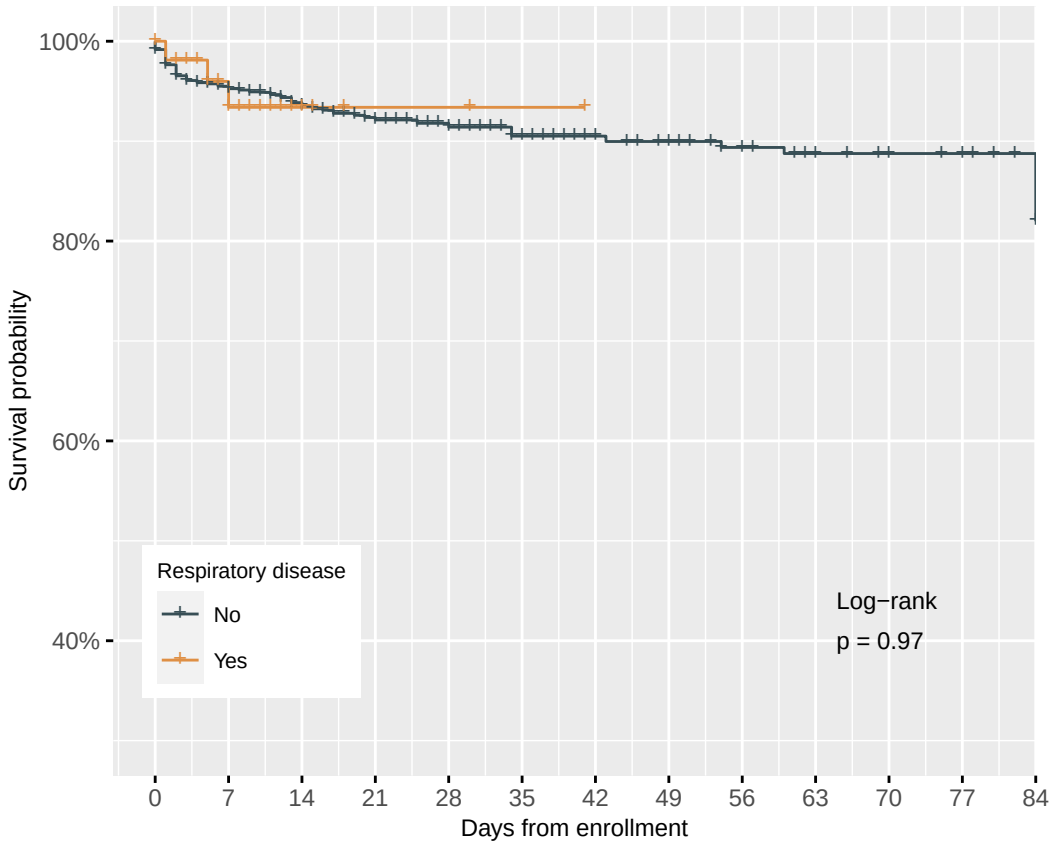

Figure 10: Survival curves of patients stratified by respiratory disease status

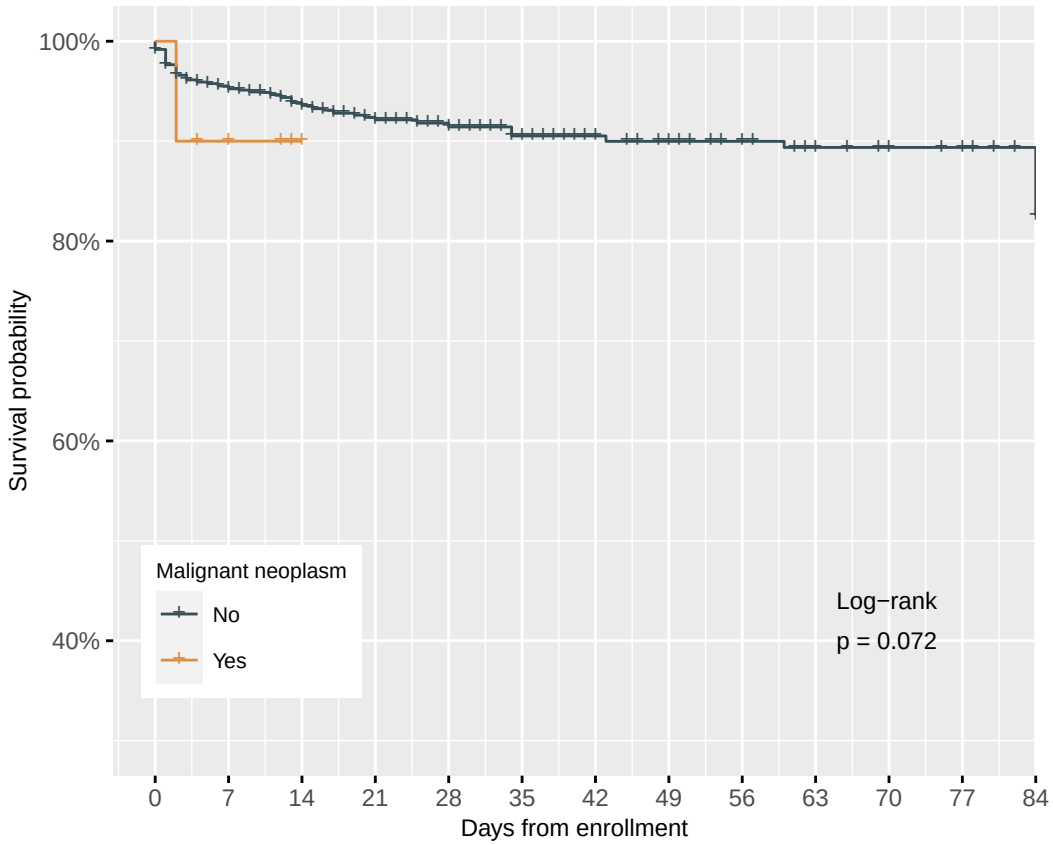

Figure 11: Survival curves of patients stratified by malignant neoplasm status

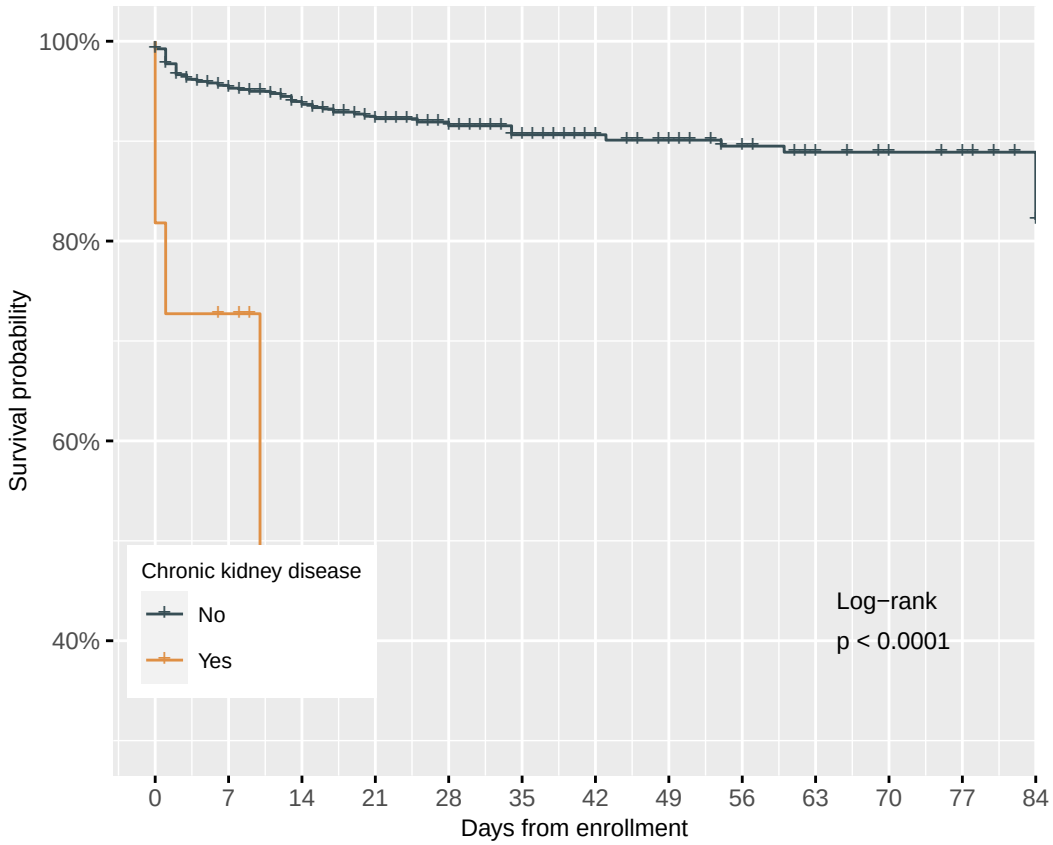

Figure 12: Survival curves of patients stratified by chronic kidney disease status

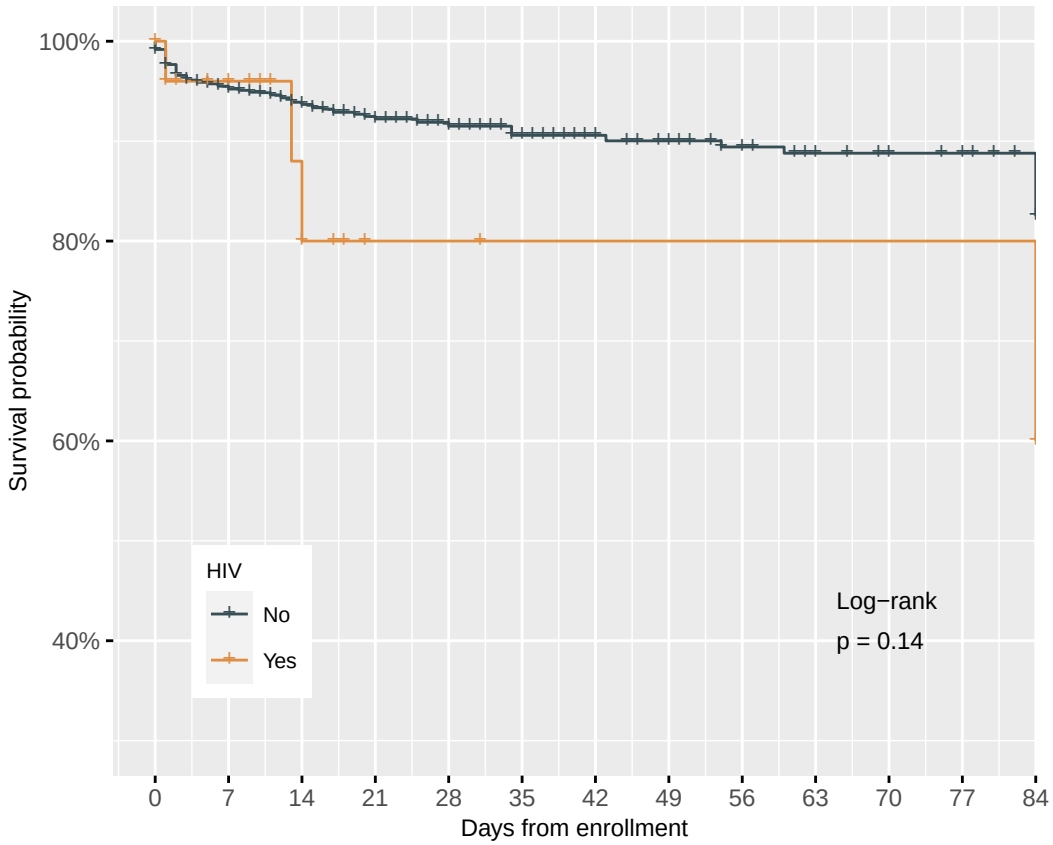

Figure 13: Survival curves of patients stratified by HIV status

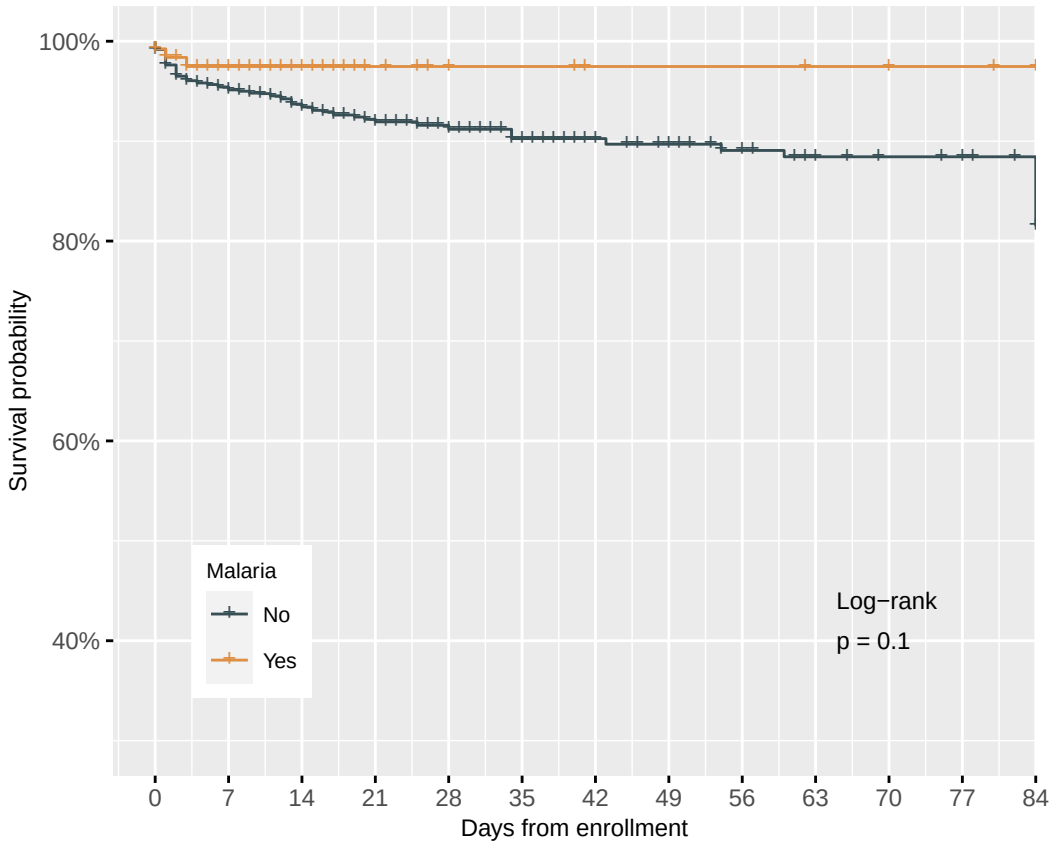

Figure 14: Survival curves of patients stratified by malaria status

Medications

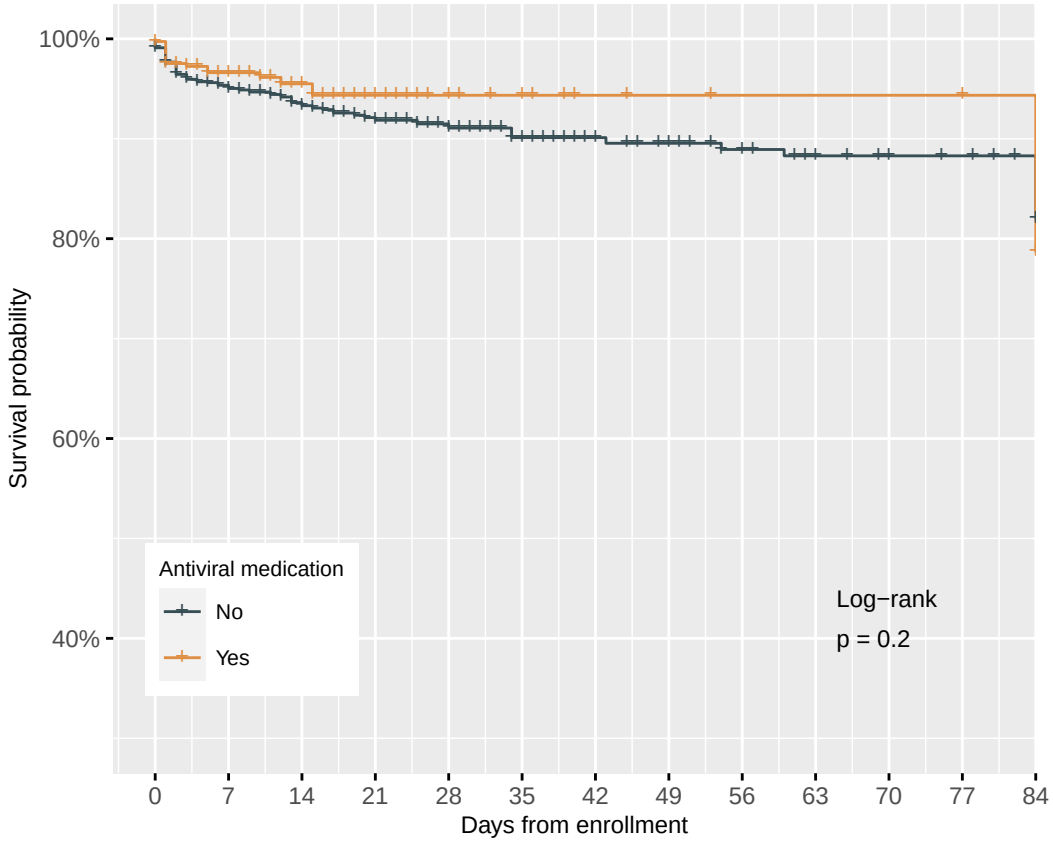

Figure 15: Survival curves of patients stratified by antiviral medication

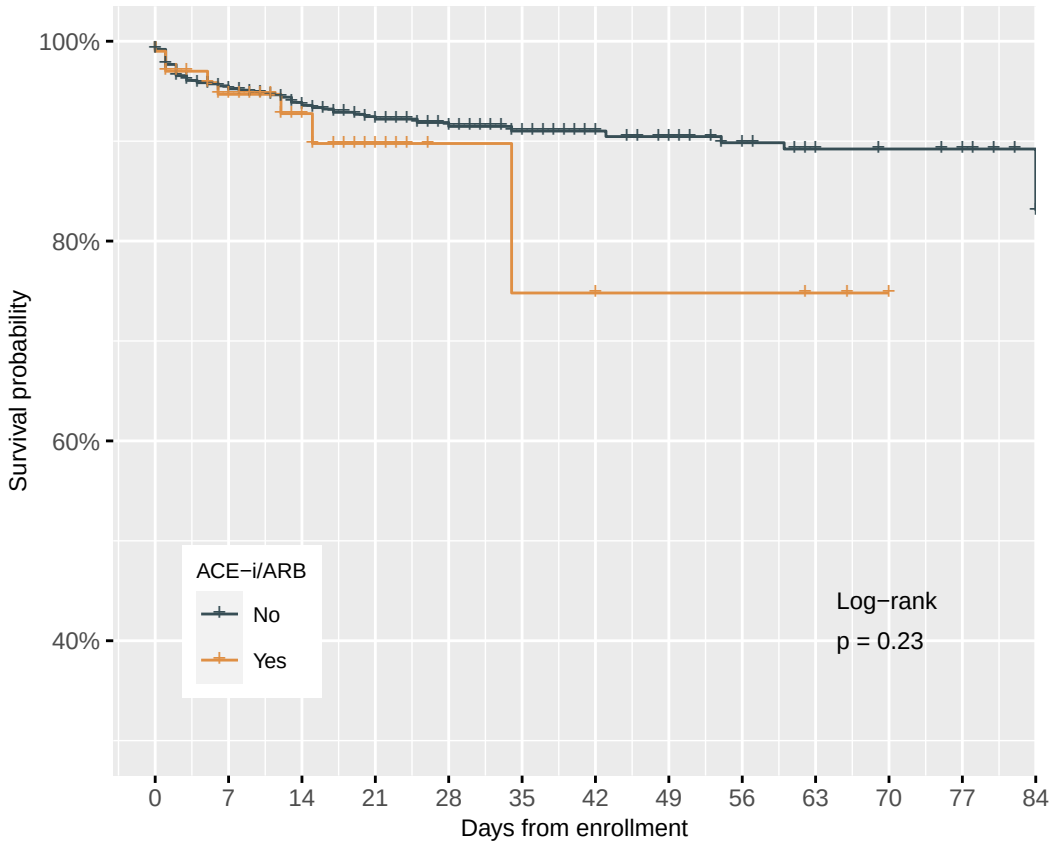

Figure 16: Survival curves of patients stratified by ACE-i/ARB medication

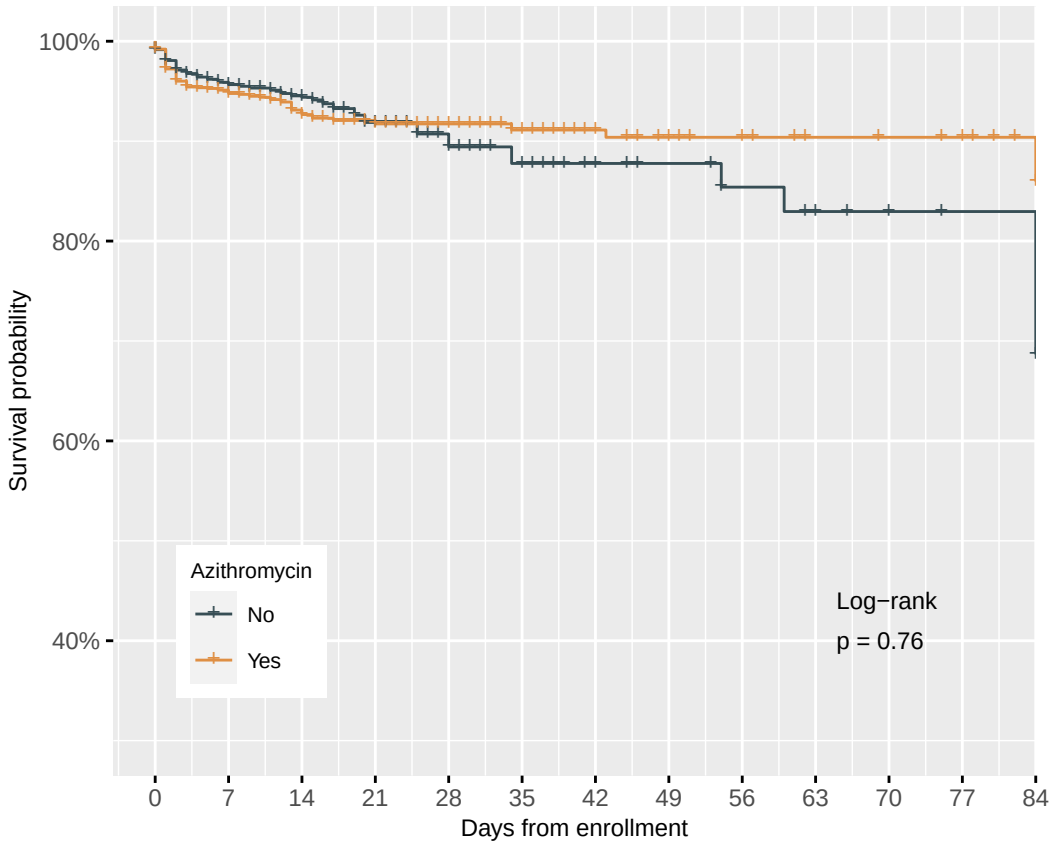

Figure 17: Survival curves of patients stratified by azithromycin use

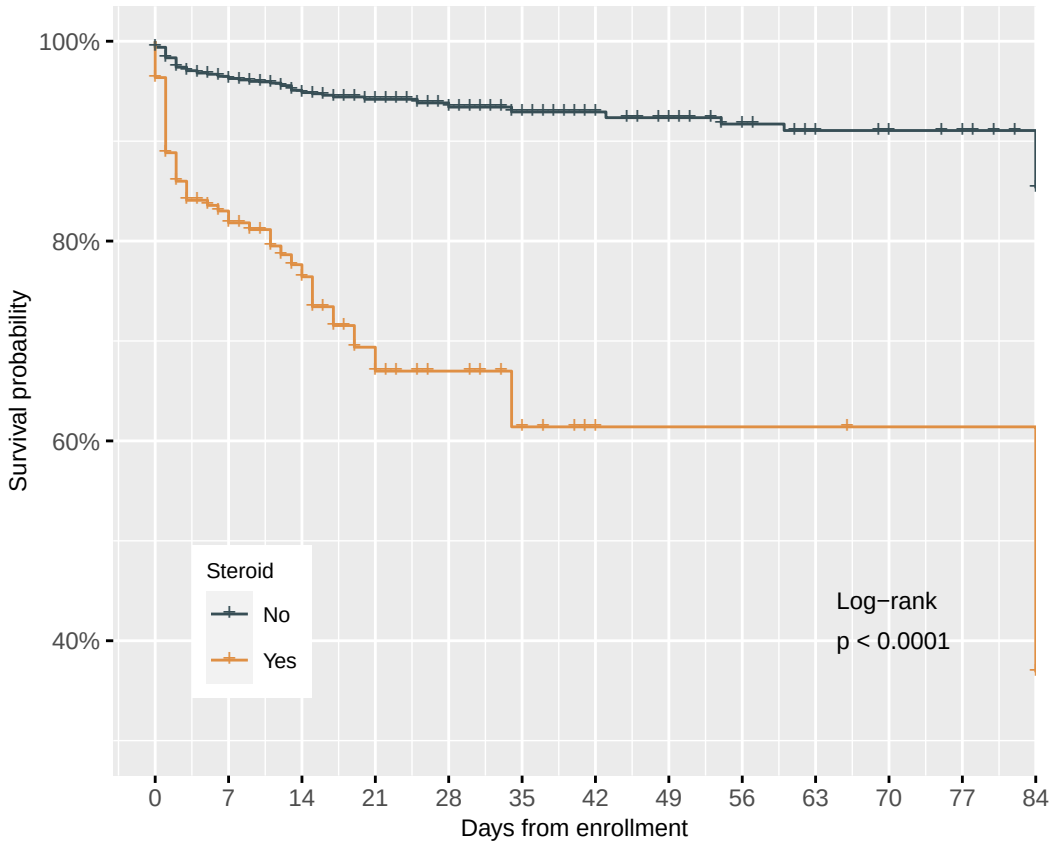

Figure 18: Survival curves of patients stratified by steroid use

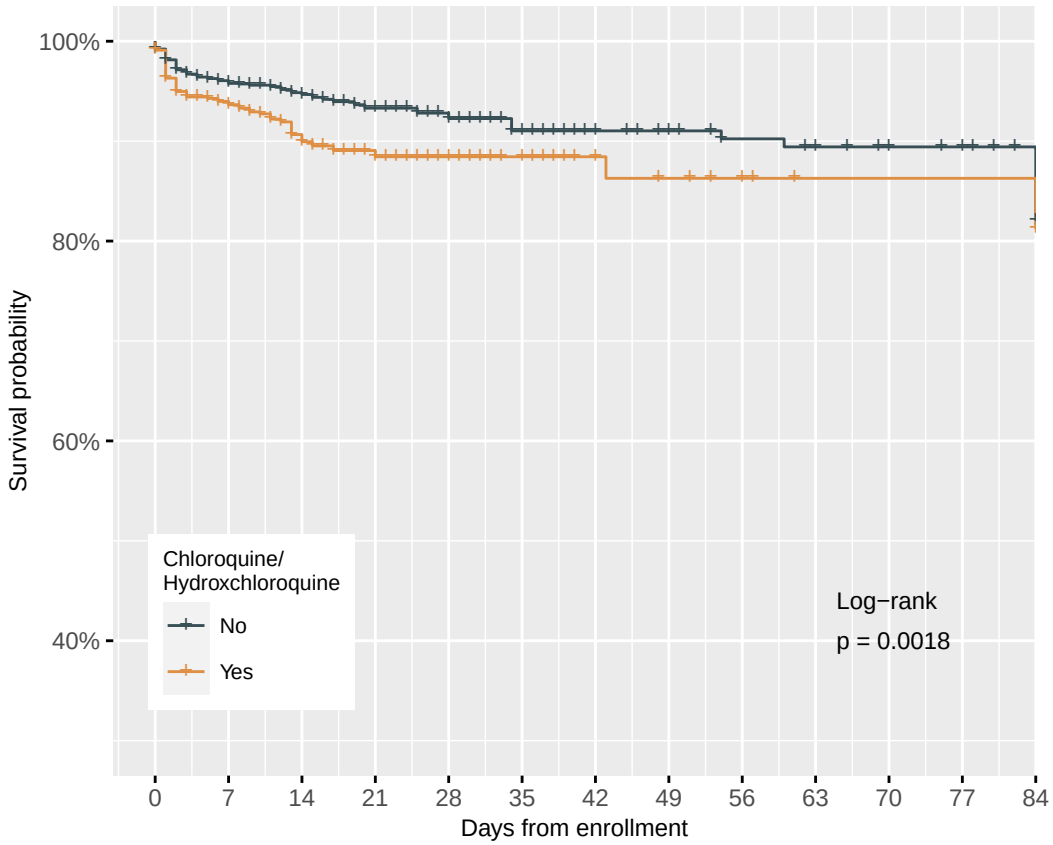

Figure 19: Survival curves of patients stratified by chloroquine/hydroxychloroquine use

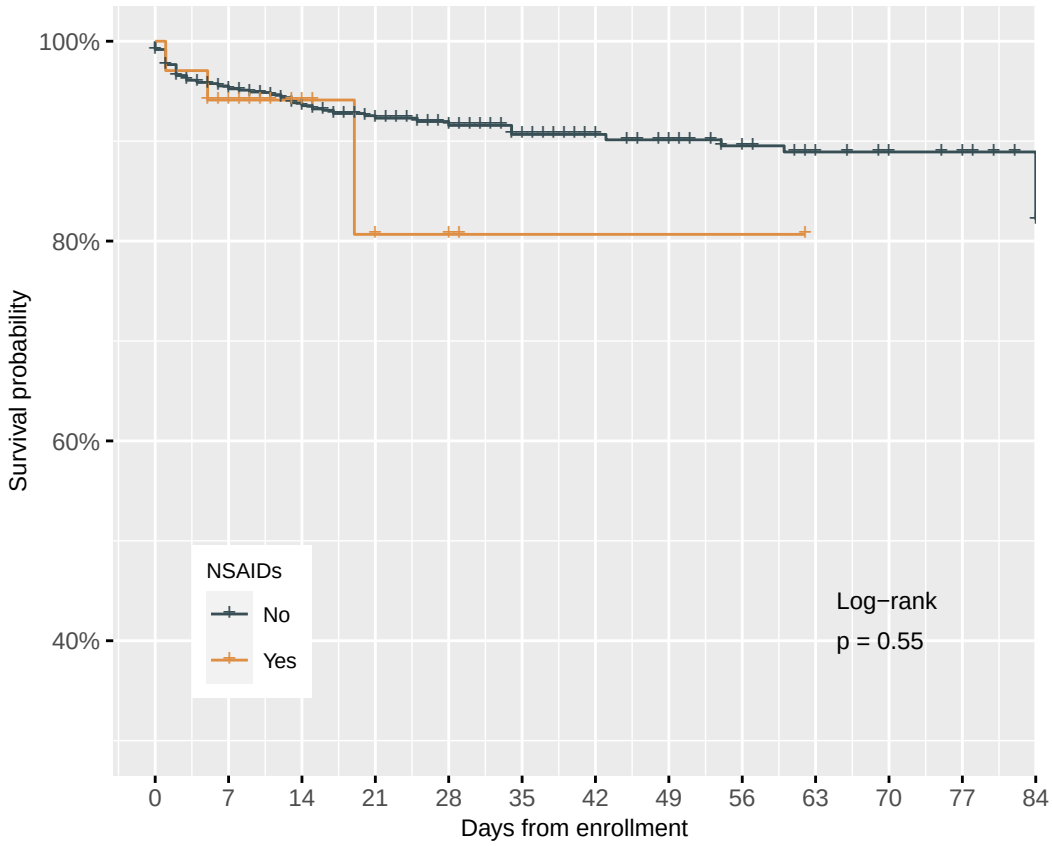

Figure 20: Survival curves of patients stratified by NSAIDs use

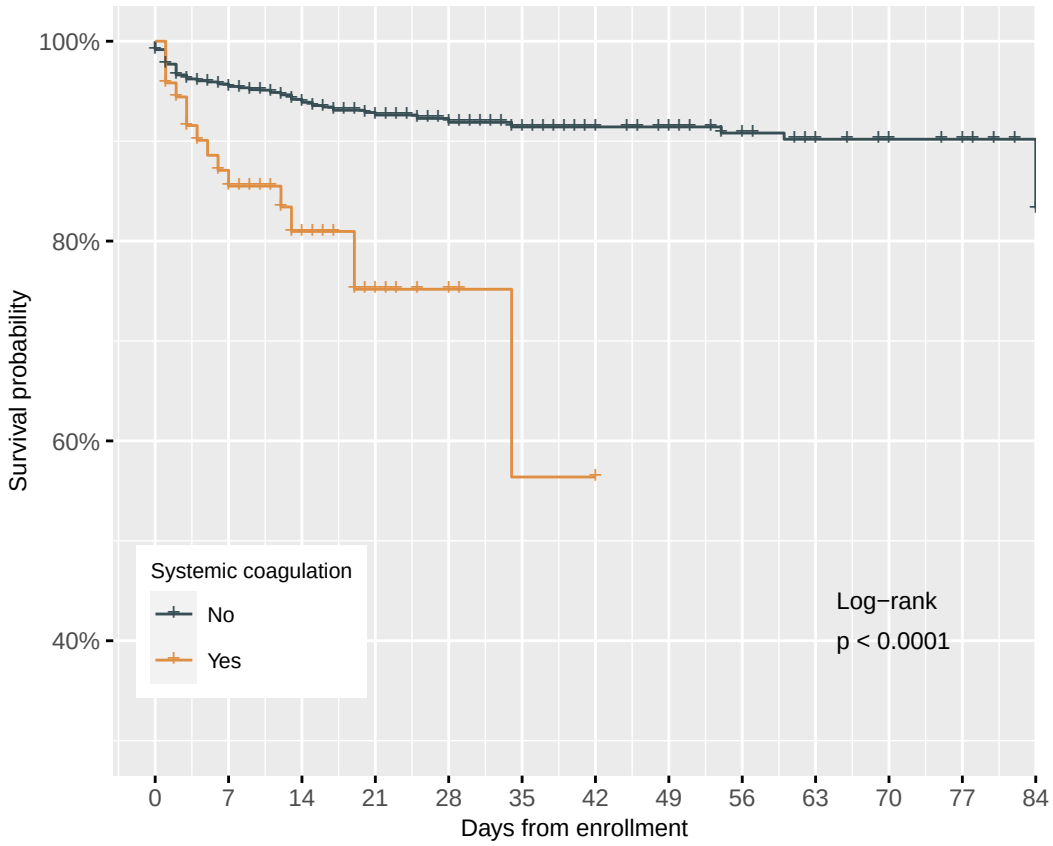

Figure 21: Survival curves of patients stratified by systemic coagulation status

Logistic regression model

We fitted a mixed effects logistic regression model using maximum likelihood and the bound optimization by quadratic approximation algorithm to predict mortality with severity, age group, male gender, health worker status, chronic cardiac disease, diabetes, respiratory disease ((defined as presence of any of chronic pulmonary disease, active or previous tuberculosis, and asthma), chronic kidney disease, malaria, azithromycin and chloroquine entered into the model as fixed effects. The model included site and facility type as random effects. The most parsimonious model arrived at through backward stepwise regression is shown in Table 2 below.

The model's total explanatory power is substantial (conditional R^2 = 0.55) and the part related to the fixed effects alone (marginal R^2) is of 0.25. The intra-class correlation coefficient (ICC) = 0.40. The conditional R^2 is the proportion of overall variance explained by the model when the site and facility type random effects are taken into account. The marginal R^2 is the proportion of the overall variance explained by the fixed effects only. The ICC is a measure of the correlation between patients within clusters. It ranges from 0 to 1. Confidence Intervals (CIs) at the 95% level and p-values were computed using the Wald approximation.

Table 2: Mixed effects logistic regression model of predictors of mortality

Predictors	**OR**	**95% CI**	*p* value
Severity score	3.50	2.73, 4.49	<0.001
Age group			
<18	...	...	
18—45	1.81	0.58, 5.61	0.307
46—65	3.93	1.29, 12.02	0.016
66—75	5.37	1.68, 17.14	0.005
>75	6.81	2.04, 22.81	0.002
Male	1.78	1.23, 2.56	0.002
Health care worker	0.22	0.06, 0.79	0.020
Chronic cardiac disease (not hypertension)	3.07	1.20, 7.86	0.019
Diabetes	2.16	1.41, 3.31	<0.001
Respiratory disease	0.34	0.09, 1.20	0.094
Chronic kidney disease	11.01	2.74, 44.25	<0.001
Malaria	0.45	0.16, 1.28	0.134
Azithromycin	0.33	0.19, 0.58	<0.001
CQ/HCQ	0.07	0.03, 0.14	<0.001
Azithromycin * CQ/HCQ	9.69	4.08, 23.04	<0.001
Number of observations	3313		
Conditional R^2	0.55		
Marginal R^2	0.25		
Log-likelihood	−507.31		
AIC	1050.61		

OR = Odds Ratio, CI = Confidence Interval, CQ/HCQ = Chloroquine/hydroxychloroquine.

The mixed effects logistic regression model showed that severity score, age, male sex, chronic cardiac disease, diabetes, and chronic kidney disease were statistically significant predictors of mortality in these hospitalized patients. Healthcare worker status, use of azithromycin, and chloroquine/hydroxychloroquine use were predictors of decreased mortality (see Table 2 and Figure 22). The forest plot shows the standardized model estimates, thus we see that age > 75 years has the largest effect on mortality while use of chloroquine/hydroxychloroquine has the largest protective effect.

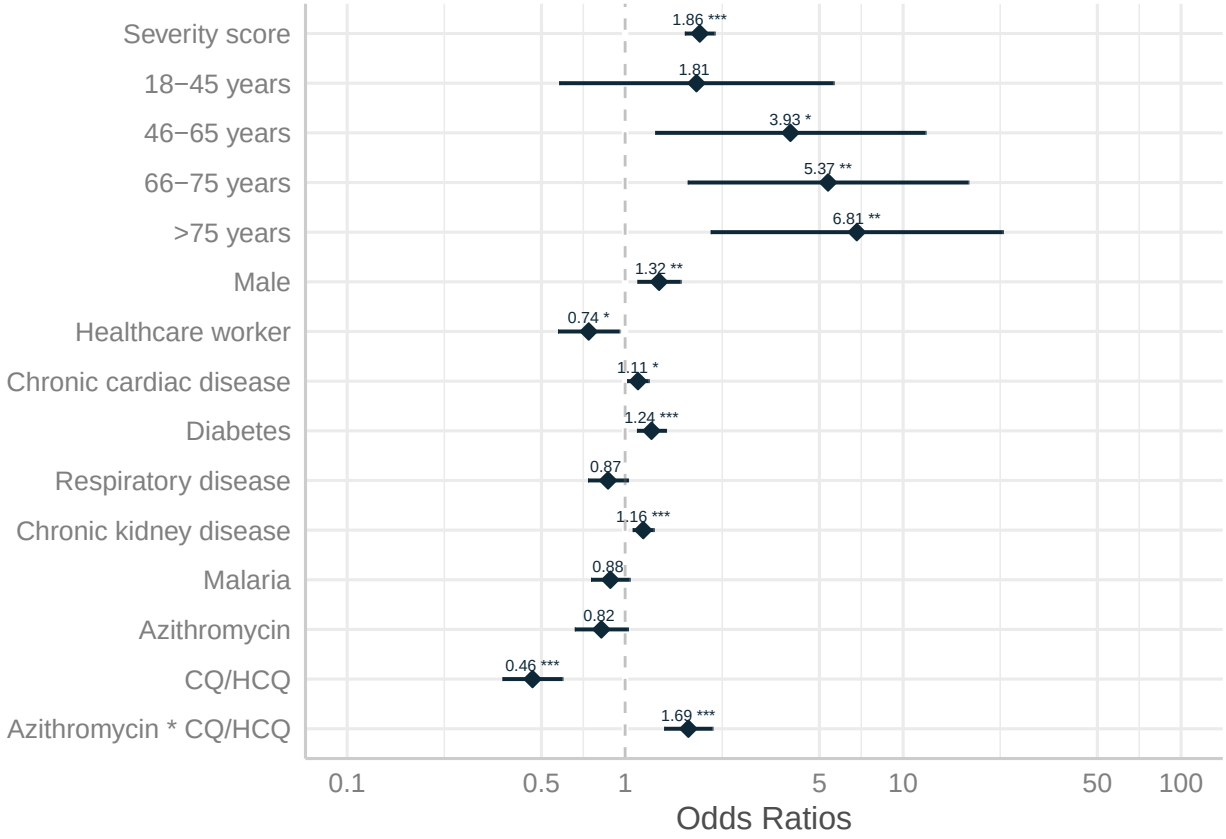

Figure 22: Forest plot of standardized model coefficients

There is an interaction effect of azithromycin and chloroquine. There is an interaction of azithromycin on the effect of chloroquine (see Figure 23). There appears to be a blunting of the effect of chloroquine on mortality by azithromycin. The predicted probability of mortality in those who did not receive chloroquine/hydroxychloroquine or azithromycin was 1.9% compared to 0.6% for those who received only azithromycin. This contrasted with a predicted probability of mortality in those who received only chloroquine/hydroxychloroquine of 0.1% compared to 0.4% for those who received both chloroquine/hydroxychloroquine and azithromycin. In other words, taking only chloroquine/hydroxychloroquine was associated with a lower probability of mortality than taking both chloroquine/hydroxychloroquine and azithromycin despite there being an association with lower mortality with both drugs alone. However, it should be noted that the size of this interaction effect is small and unlikely to be clinically meaningful. It may not be reproducible and it may be a spurious finding arising as a result of unknown patient or study site characteristics, although we tried to account for site effects by including site as a random effect in the regression model.

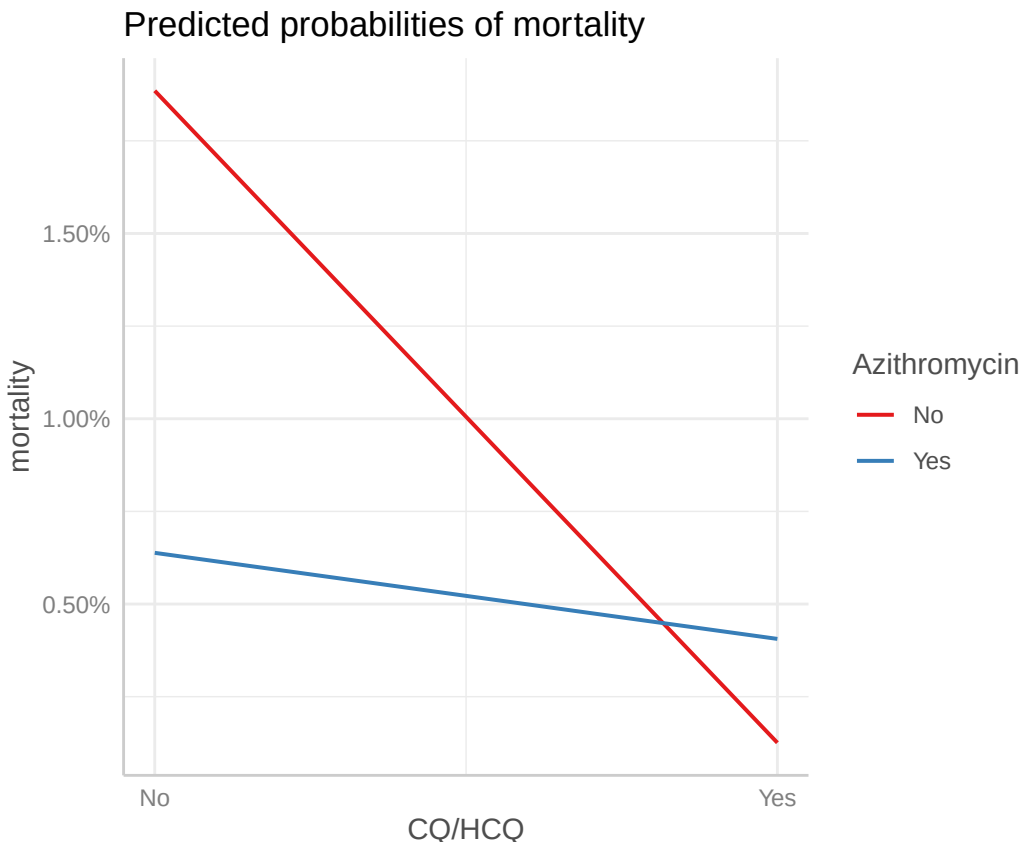

Figure 23: Interaction between azithromycin and chloroquine

Cox regression model

Table 3: Mixed effects cox regression model of predictors of time to mortality

Predictors	HR	95% CI	p value
Severity score	2.57	2.11, 3.12	<0.001
Age group			
<18	...	...	
18—45	1.47	0.44, 4.95	0.531
46—65	2.66	0.82, 8.66	0.104
66—75	3.54	1.06, 11.83	0.040
>75	3.58	1.05, 12.15	0.041
Male	1.66	1.19, 2.32	0.003
Health care worker	0.35	0.11, 1.13	0.079
Chronic cardiac disease (not hypertension)	2.34	1.22, 4.48	0.010
Diabetes	1.99	1.41, 2.80	<0.001
Respiratory disease	0.29	0.09, 0.94	0.039
Chronic kidney disease	8.21	2.86, 23.56	<0.001
Malaria	0.42	0.15, 1.20	0.104
Azithromycin	0.55	0.35, 0.86	0.008
CQ/HCQ	0.20	0.10, 0.38	<0.001
Azithromycin * CQ/HCQ	4.57	2.12, 9.85	<0.001
Number of observations	3178		
Number of events	192		
site σ^2	0.56		
facility type σ^2	1.51		
Log-likelihood	—1156.94		
AIC	2367.52		

HR = Hazard ratio, CI = Confidence Interval

Table 3 shows the results of the Cox regression model with site and facility type included in the model as shared frailty terms to account for their random effects. Table 3 presents similar results to those of the mixed effects logistic regression model shown in table 2. Statistically significant increased hazards for mortality are seen for increasing severity score, increasing age, male sex, chronic cardiac disease, diabetes, and chronic kidney disease in these hospitalized patients. Healthcare worker status, use of azithromycin, and chloroquine/hydroxychloroquine use were associated with reduced hazard ratios. Respiratory disease seems to be associated with reduced hazards ratio as well.

References

1. Hsieh FY, Bloch DA, Larsen MD. A simple method of sample size calculation for linear and logistic regression. Statistics in Medicine. 1998;17(14):1623–34.
2. Faul F, Erdfelder E, Buchner A, Lang A-G. Statistical power analyses using G*Power 3.1: Tests for correlation and regression analyses. Behavior Research Methods. 2009 Nov;41(4):1149–60.
3. Subramanian SV. Multilevel Methods, Theory, and Analysis. In: Encyclopedia of Health and Behavior. 2455 Teller Road, Thousand Oaks California 91320 United States: SAGE Publications, Inc.; 2004.
4. Gelman A, Hill J. Data analysis using regression and multilevel/hierarchical models. Cambridge; New York: Cambridge University Press; 2007.
5. R Core Team. R: A Language and Environment for Statistical Computing. Vienna, Austria: R Foundation for Statistical Computing; 2021.
6. Therneau TM. Coxme: Mixed Effects Cox Models. 2020.