ORIGINAL ARTICLE

Effect of COVID-19 Vaccination on Survival of Iranian Patients: comparison of Cox Regression, Survival Tree and Forests

Mozhgan Seif[1], PhD; Mehdi Sharafi[2], PhD; Marziyeh Sadat Seif[3], MSc; Mohsen Bayati[4], PhD; Jalal Karimi[5], MD

[1]Non-communicable Diseases Research Center, Department of Epidemiology, School of Health, Shiraz University of Medical Sciences, Shiraz, Iran
[2]Department of Community Medicine, School of Medicine, Hormozgan University of Medical Sciences, Bandar Abbas, Iran
[3]Seyed-al-Shohada Hospital, Jahrom University of Medical Sciences, Jahrom, Iran
[4]Health Human Resources Research Center, School of Health Management and Information Sciences, Shiraz University of Medical Sciences, Shiraz, Iran
[5]Department of Infectious Diseases, School of Medicine, Fasa University of Medical Science, Fasa, Iran

Correspondence:
Mehdi Sharafi, PhD;
Department of Community Medicine, School of Medicine, Hormozgan University of Medical Sciences, Bandar Abbas, Iran
Email: shatafi_2002@yahoo.com
Mohsen Bayati, PhD;
Health Human Resources Research Center, School of Health Management and Information Sciences, Shiraz University of Medical Sciences, Shiraz, Iran
Email: bayatim66@gmail.com
Received: 04 April 2023
Revised: 25 May 2023
Accepted: 17 June 2023

Abstract

Background: The effectiveness of COVID-19 vaccination determines the resource allocation for saving lives in this pandemic. Certainly, the efficacy of all vaccines has been studied in laboratory situations. However, the present study aims to estimate the effectiveness of vaccination in real conditions of Iranian populations by controlling the effect of demographic factors and the history of chronic diseases.

Methods: This historical cohort study used information on 1988 hospitalized COVID-19 patients with less than 93% blood oxygen levels. Cox regression, Survival Tree, and Forests were applied to estimate the effect of immunity from vaccination on survival, while the effect of demographic characteristics and history of chronic diseases was controlled.

Results: The analysis showed that 10% of the patients were immunized, while the Sinopharm vaccine and the rest by Sputnik induced 86% of the immunity. Although there was no difference in the effectiveness of the vaccines, it was found that the immunity from each vaccine increased survival. Patient age was identified as the most influential factor in survival. Other contributing factors include the history of opium/smoking, cancer, the history of chronic lung disease, kidney disease, high blood pressure, and cardiovascular problems.

Conclusion: Although age is the most influential factor in patient survival, immunity can control the risk of COVID-19. Therefore, it is recommended to prioritize the old and patients with any of the above underlying problems for vaccination and timely treatment.

Please cite this article as: Seif M, Sharafi M, Seif MS, Bayati M, Karimi J. Effect of COVID-19 Vaccination on Survival of Iranian Patients: comparison of Cox Regression, Survival Tree and Forests. J Health Sci Surveillance Sys. 2023;11(Supplement 3):620-630.

Keywords: Cog regression, COVID-19, Vaccination, Survival analysis, Survival tree and forests

Introduction

The pandemic disease SARS-COV2 (COVID-19) is still rapidly intensifying worldwide.[1, 2] According to the World Health Organization (WHO), Globally, as of 17 August 2022, there have been 589,680,368 confirmed cases of COVID-19, including 6,436,519 deaths, reported to WHO. As of 8 August 2022, a total of 12,355,390,461 vaccine doses have been administered.[3] In Iran, from 3 January 2020 to 17 August 2022, there have been 7,488,493 confirmed cases of COVID-19 with 143,093 deaths reported to WHO. As of 26 July 2022, 152,446,648 vaccine doses have been administered.[2] Common symptoms of COVID-19 include fever, cough, fatigue, shortness of breath, muscle aches, and other symptoms, including sore throat, chest tightness, nausea, vomiting, diarrhea, headache, and so on . Also, it must be noted that about 80% of patients have no symptoms or mild disease in a way that can be cured only by self-medication.[4, 5] Most disease transmission occurs at the onset of the disease in the infected person, although the disease is transmitted before the onset of symptoms.[6] So far five variants of COVID-19, including Alpha (Date of designation-Dec-2020), Beta (18-Dec-2020), Gamma (11-Jan-2021), Delta (11-May-2021), and Omicron (26-Nov-2021) have been identified. Despite global efforts to develop vaccines and widespread vaccinations, viral mutations and the emergence of new virus strains, such as the recent strain of Omicron, have negatively affected this viral disease control.[7]

Studies around the world have shown that there are several risk factors, including high blood pressure and old age, male gender, and comorbidities such as diabetes, chronic lung disease, heart, liver, kidney problems, tumors, immunodeficiency, and pregnancy, which affect the exacerbation of COVID-19 disease.[8-13] In other studies, with the Random Survival Forest algorithm, factors like old age, cardiovascular disease,[14] respiratory distress, and kidney problems[15, 16] are known to be important risk factors. In another study by Survival Tree, age over 65 years and female gender were identified as important factors in patient survival.[17]

Previous clinical trials have reported a positive effect of COVID-19 vaccination on survival, reduction of severe disease, and transmission to others.[18-20] What seems important is the need for vaccination to manage the pandemic.[20] However, the effectiveness of vaccines in clinical trial studies may differ from their effectiveness in community and field studies.[21] Therefore, further observational studies are needed to investigate the effect of vaccination on the field and the actual conditions of communities. Therefore, the present study aims to determine the effect of vaccination on the survival of Iranian COVID-19 patients. Multivariate models and algorithms are often used to provide a more accurate estimation and control for confounding effects, especially when health predictors such as COVID-19 symptoms and comorbidities are included in the models. Likewise, to control the confounders, an appropriate statistical model is required for accurate estimation of vaccines' effects. Therefore, three different models, including Cox regression, Survival Tree, and Random Survival Forest, are compared in this study. It is worth noting that no study has used conditional survival trees and random survival forests to estimate vaccination efficacy.

Methods

Study Population

This study was performed as a Historical Cohort on hospitalized patients with COVID-19 symptoms referred to Vali Asr Hospital of Fasa, southern Iran. From March 11, 2020, to December 11, 2021, 6543 (87%) out of 7511 patients with COVID-19 symptoms were referred to this hospital, with positive PCR test result. Among patients with COVID-19 positive tests, patients with blood oxygen saturation levels less than 93% were 2641 (40%). From this final set, 1988 patients with complete information about the length of hospital stay, patient's final status (death/discharge), and vaccination status, were included in the analysis. The hospitalization time was considered the starting point of the follow-up; and the patients were followed until death/discharge. In this study, individuals whose vaccination was done two weeks before considered fully vaccinated. It is worth mentioning that the entire population was vaccinated with only two types of vaccines, including Sinopharm and Sputnik V.

Inclusion criteria: All patients hospitalized in Vali Asr Fasa Hospital due to Coronavirus (COVID-19) from March 11, 2020, to December 11, 2021, were included in the study.

Exclusion criteria: patients whose information (status: death/discharge, duration of hospitalization) was not recorded were excluded from the study. Also, the variables with more than 20% missing were removed.

Measured Variables

The variables used in this study are immunity from vaccination (or fully vaccinated), age, history of opium/smoking, cancer, history of chronic lung disease, kidney disease, high blood pressure, cardiovascular problems, chronic liver disease, blood problems, pregnancy, diabetes history, gender, length of hospital stay, and date of death or discharge.

Survival Tree Algorithm

Survival Tree is a common type of decision tree. Growth in this tree takes place in such a way that patients with the same survival probability would be classified in the same nodes. According to the log-Rank test, if patients' survival is at different variable levels, that variable is considered a split variable or tree branch, and the tree grows based on it. Therefore, tree growth is based on variables that affect survival. In other words, tree branches can be considered variables affecting survival. Indeed, the effect of each variable in the tree algorithm is estimated by conditional inference. To control family-wise error due to applying multiple tests, the authors used Bonferroni adjustment.[22]

Random Survival Forest Algorithm

It should be noted that an ensemble of trees forms a Forest. Each tree in the forest is formed based on Bootstrap sampling. The data used to grow the trees are varied because 37% of the data in each Boot strap sample is duplicated, so different trees are formed in a forest. Forest estimates have less variance than the trees'. Therefore, the forest is a good algorithm for estimating the importance of variables. To estimate the importance of each variable, the forest is firstly grown with the actual data and then with the permuted data on the desired variable. The difference in the accuracy of predictions in these two conditions determines the importance of the variable.

Statistical Analysis

To analyze, the first variables with more than 20% missing values were excluded.[23] The remaining missing values were then imputed using a random survival forest algorithm. In addition to t-test, Chi-square and Fisher's exact test, the analytical methods used included a survival tree and forest consisting of 1000 trees. The tree was pruned to prevent over-fitting. Furthermore, the Cox regression has been the most widely used regression in survival analysis.[22] Indeed, stepwise variable selection, according to Akaike Information Criterion (AIC), is used to finalize Cox regression. Stepwise variable selection is a repetitious algorithm in that the addition/removal of every variable is assessed in consecutive steps. Also, Proportional Hazard assumption, a preconception assumption for fitting Cox regression is assessed in the analysis.

It is worth noting that to estimate the effect of vaccination immunity accurately, other demographic variables and comorbidities have been adjusted in all three models. The variables used in the models are age, history of opium/cigar use, cancer, immunity, history of chronic lung disease, kidney disease, high blood pressure, cardiovascular problems, chronic liver disease, blood problems, pregnancy, history of diabetes, and gender.

Harrell's Concordance (C) index was used to compare the performance of random survival forest and Cox regression. Finally, P<0.05 was considered statistically significant, and all the analysis was performed in R (R Core Team (2020). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/.)version 3.1.0, a freely available software. The packages mainly used include openxlsx, party, randomForestSRC, and survival.

Results

Data from 1988 hospitalized COVID-19 patients with less than 93% blood oxygen levels showed that 386 (19%) patients died. The median duration of hospitalization was seven days, and the mean blood oxygen level was 86%±9%. Due to vaccination, 10% of patients were immunized. The death rate in the immune patients was 12% and 20% in the others (P=0.008). Notably, 86% of the injected vaccines were Sinopharm, and the rest were Sputnik V, but no differences were observed in hazard reduction for these two vaccines (P=0.90). Death hazard for patients receiving Sinopharm and Sputnik V vaccines

Table 1: Demographic Variables of Hospitalized Covid-19 Patients with Blood Oxygen Levels Less than 93%

Variables	**Group**	**Final Statue of Covid-19 Patients**		**P value**
		Dead N (%)	**Survived N (%)**	
Sex	Male	180 (46.63)	710 (44.32)	0.412[a]
	Female	206 (53.37)	892 (55.68)	
Vaccination Immunity	No	361 (93.52)	1426 (89.01)	0.008[a]
	Yes	25 (6.48)	176 (10.99)	
Hist. Chronic Pulmonary Dis.	No	368 (95.34)	1591 (99.31)	<0.001[b]
	Yes	18 (4.66)	11 (0.69)	
Hist. Chronic Kidney Dis.	No	374 (96.89)	1597 (99.69)	<0.001[b]
	Yes	12 (3.11)	5 (0.31)	
Hist. Hypertension	No	213 (55.18)	1206 (75.28)	<0.001[a]
	Yes	173 (44.82)	396 (24.72)	
Hist. Cardiovascular Dis.	No	318 (82.38)	1491 (93.07)	<0.001[a]
	Yes	68 (17.62)	111 (6.93)	
Hist. Diabetes	No	285 (73.83)	1347 (84.08)	<0.001[a]
	Yes	101 (26.17)	255 (15.92)	
Hist. Chronic liver Dis.	No	384 (99.48)	1599 (99.81)	0.251[b]
	Yes	2 (0.52)	3 (0.19)	
Cancer	No	379 (98.19)	1591(99.31)	0.036[a]
	Yes	7 (1.81)	11 (0.68)	
Opium/Cigar smoking	No	380 (98.44)	1595 (99.56)	0.026[b]
	Yes	6 (1.55)	7 (0.44)	
Pregnancy	No	385 (99.74)	1594 (99.50)	0.999[b]
	Yes	1 (0.26)	8 (0.50)	
Blood Dis.	No	384 (99.48)	1590 (99.25)	0.999[b]
	Yes	2 (0.52)	12 (0.75)	
Age (years)		Mean±SD	Mean±SD	<0.001[c]
		67.49±15.97	54.30±16.58	

was 63% and 55% lower, respectively (all P<0.05). The mean age in the whole population was 57±18 years. The mean age of expired patients (67±16 years) was higher than the mean age of discharged ones (54±17 years) (P<0.05). Table 1 presents the comparison of discharged and expired patients. This comparison showed that the history of opium/cigar smoking, cancer, chronic lung disease, kidney disease, high blood pressure, cardiovascular problems, and diabetes in expired patients was more than in discharged patients (all P<0.05).

Kaplan–Meier Plot

Figure 1 shows the Kaplan–Meier survival curve of the whole population. Accordingly, the cure rate was 65%. Survival rates at 7, 14, and 21 days were 88%, 78%, and 70%, respectively. After 21 days, the plot's slope was very smooth, and there wasno significant decline in the plot. Figure 2 shows patients' Kaplan–Meier survival curve according to their immune status. As can be seen in the figure, the survival of immune patients is higher than that of the other patients. Notably, the Log-Rank test also confirmed this difference (P<0.05).

Survival Tree

Figure 3 illustrates the survival tree. The first variable on which the tree grew was the age of the patients. Therefore, according to the tree algorithm, age can be considered the most important covariate in determining patients' survival. After the age of patients, the tree grew based on the variables of immunity, history of hypertension, and cardiovascular problems.

Random Survival Forest

It should be noted that the tree's all-recognized

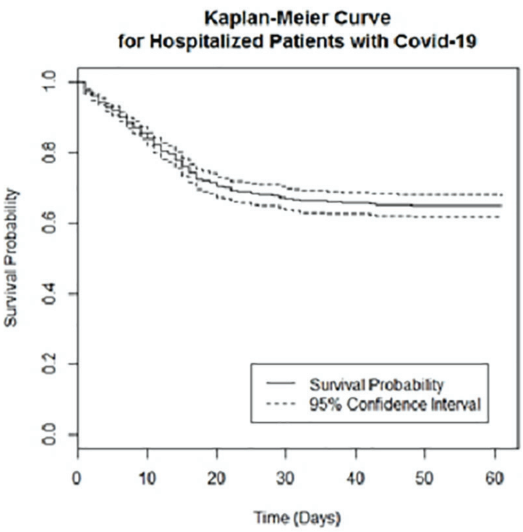

Figure 1: The Kaplan–Meier survival curve of the whole population of patients

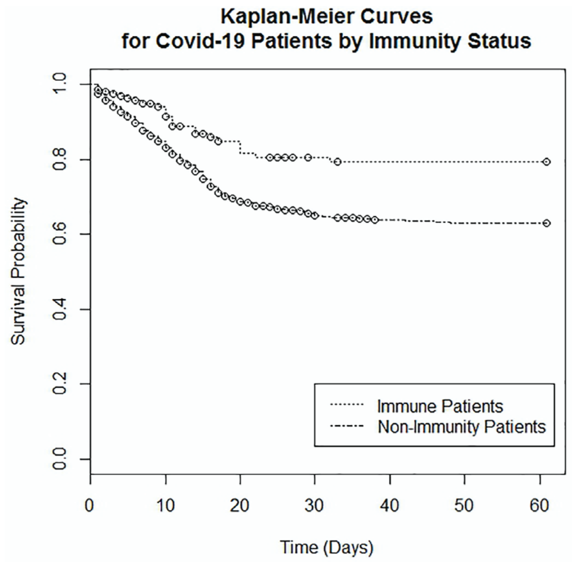

Figure 2: The Kaplan–Meier survival curve of patients according to their immune status.

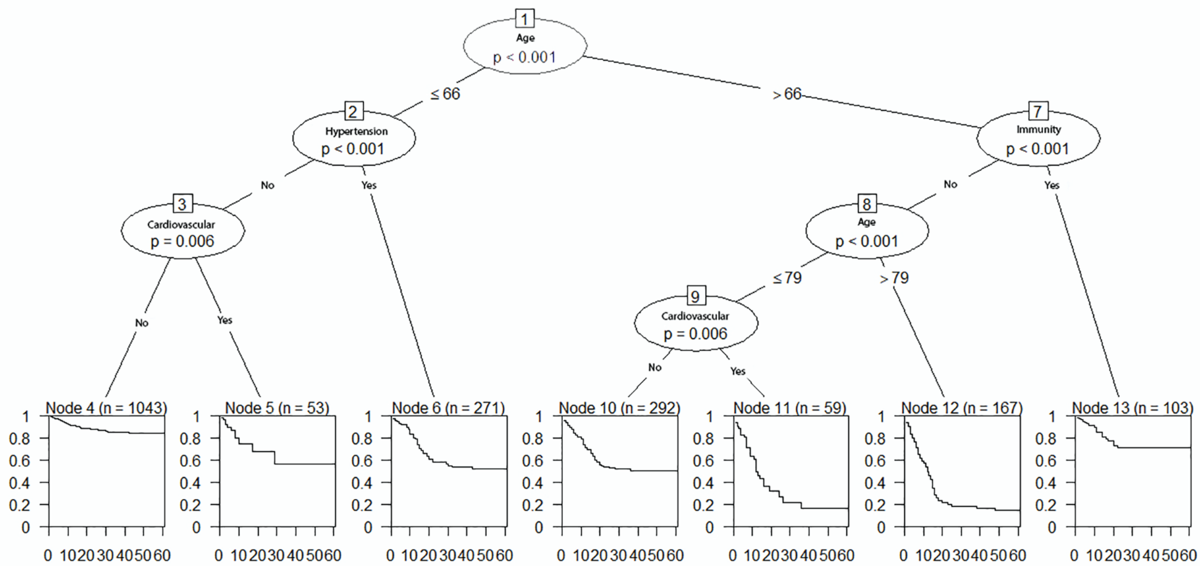

Figure 3: Conditional survival tree of hospitalized COVID-19 patients with blood oxygen level less than 93%

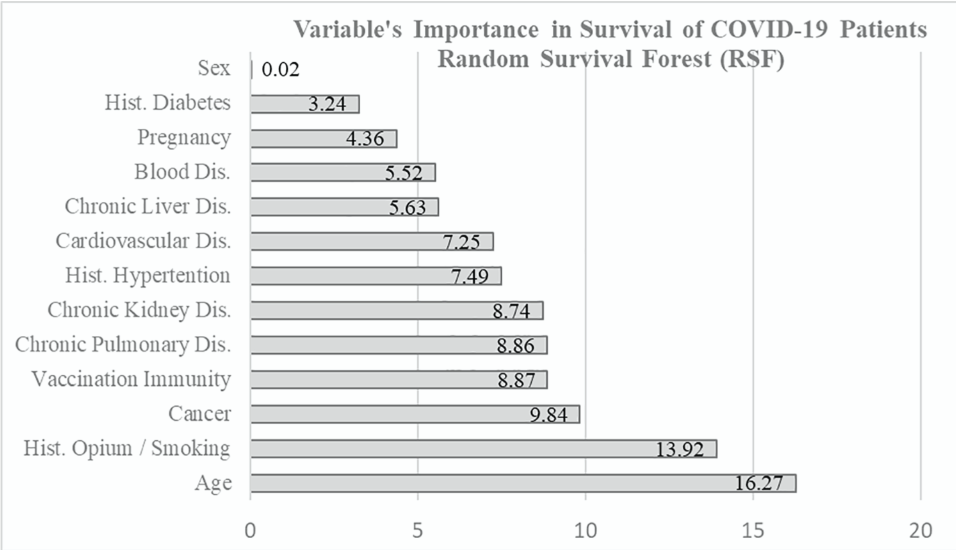

Figure 4: Variable's importance in survival of COVID-19 hospitalized patients by Random Survival Forest (RSF) algorithm.

Table 2: Cox regression for modelling the survival of COVID-19 patients with low blood oxygen level at the time of admission to the Vali-Asr Fasa Hospital

Factor	Coefficient	S.E.	HR*	95%CI for HR		P value	P value of PH**
				Lower	Upper		
Age	0.038	0.003	1.039	1.032	1.046	<0.001	0.982
Opium/Cigar	1.793	0.434	6.009	2.564	14.085	<0.001	0.527
Cancer	1.098	0.384	2.998	1.413	6.363	<0.001	0.127
Vaccination Immunity	-1.176	0.210	0.308	0.204	0.465	<0.001	0.940
Hist. Chronic Pulmonary Dis.	0.910	0.247	2.483	1.523	4.032	<0.001	0.081
Hist. Chronic Kidney Dis.	0.801	0.299	2.229	1.240	4.005	0.007	0.430
Hist. Hypertension	0.335	0.114	1.398	1.116	1.750	0.003	0.320
Hist. Cardiovascular Dis.	0.533	0.139	1.704	1.297	2.239	0.001	0.797
Hist. Chronic liver Dis.	1.331	0.712	3.786	0.936	15.305	0.062	0.888
Hist. Diabetes	0.231	0.124	1.260	0.988	1.609	0.062	0.067
Model						<0.001	0.248

*Hazard Ratio; **Proportional Hazard

variables, are also important in the forest algorithm. Forest respectively recognized age, history of opium use, and cancer as the most important variables in predicting survival (Figure 4). The immunity variable from vaccination is the fourth most important, followed by the history of chronic lung disease, kidney disease, high blood pressure, and cardiovascular problems. The history of chronic liver disease, blood diseases, pregnancy, history of diabetes, and gender were less important.

Stepwise Proportional Hazard Cox Regression

Table 2 presents the results of the final Cox regression. The last column of Table 2 represents the P values regarding the Proportional Hazard (PH) assumption. For all variables and the model, all P values are greater than 0.05; therefore, the proportional hazards assumption is established for each of the variables and generally accepted for the Cox regression. The variables of gender, pregnancy, and history of blood disease were removed from the model according to the step-by-step variable selection method and did not remain in the final model. The results showed that the hazard of death increases by 4% for every one year rise in age.

The history of opium/cigar use and cancer increases the hazard by 6 and 3 times, respectively. COVID-19 immunity reduces the risk of death by 70%; in other words, the lack of immunity increases the hazard of death more than three times. A history of chronic lung and kidney disease increases the hazard more than twice. A history of high blood pressure and cardiovascular disease increases the risk by 40% and 70%, respectively (all P<0.05). A history of chronic liver disease increases the hazard by 3.8 times (P=0.062). A history of diabetes increases the hazard by 26%; however, the p-value related to this factor is 0.062 and is not significant. To further analyze diabetes' effect, this variable was exclusively included in the Cox model, showing that diabetes increases the hazard of death by 85% (P<0.05). In addition, Figure 5 depicts a survival tree that is fitted only in the presence

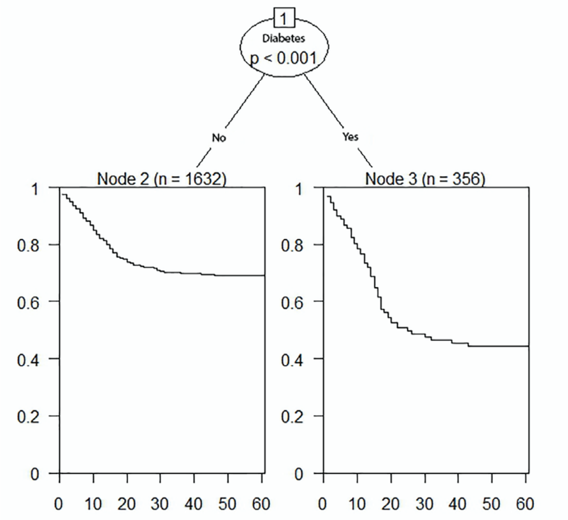

Figure 5: Survival tree of COVID-19 hospitalized patients that is fitted only in the presence of the diabetes variable

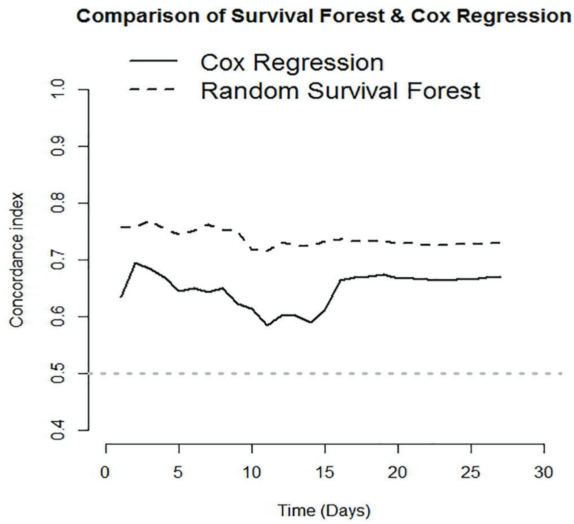

Figure 6: Concordance Index to compare Cox Regression and Random Survival Forests for modeling the survival of COVID-19 hospitalized patients with blood oxygen less than 93%

of the diabetes variable. It is clear that the tree's growth based on the diabetes variable and the split of the mother node showed the effect of diabetes on the survival of patients using this technique (P<0.05).

Specific Analysis Regarding the Effect of Diabetes

To further analyze and find out the reason for the insignificance of diabetes in multiple analyzes, the independence of diabetes and cardiovascular disease was examined, strongly rejecting the hypothesis of independence (P=0.025). Further investigation showed that the concordance between diabetes and cardiovascular disease was 77%, so 77% of patients have had the same status for both variables.

Comparison of Applied Models

Finally, Figure 6 displays the C-index to compare the performance of random forest and Cox regression. As can be seen, the Randon survival forest's C-index has been higher for the whole duration.

Discussion

The results of the present study on COVID-19 patients with blood oxygen levels less than 93% showed that the injection of both Sinopharm and Sputnik vaccines increased survival. Cox regression models, survival tree, and random survival forests identified patients' age as the most important covariate affecting patients' survival. The immunity from vaccination was another factor whose effectiveness was confirmed by all methods. Before immunity, the most important variables that affected the survival of patients were the history of opium/smoking use and cancer. In agreement with the survival forest, in the Cox regression, a history of chronic lung disease, kidney disease, high blood pressure, and cardiovascular problems were found to be hazardous. The history of liver disease, blood problems, pregnancy, history of diabetes, and gender were insignificant in the Cox model and less important in the survival forest.

As mentioned, the age of patients was determined as the most important covariate in patient survival by all methods used. Perhaps the primary cause of this problem is the changes in the immune system during old age. The ability of the immune system to control the viral load determines whether the patient becomes sick slightly or severely. This ability is reduced in old age due to changes in immune function called immunosenescence. These changes prevent pathogen detection, alert signaling, and pathogen clearance.[24-26] Therefore, the severity and outcome of COVID-19 disease largely depend on the patient's age. Other studies have found that people over the age of 65 form 80% of inpatients and are 23 times more likely to die.[24, 27] It is noteworthy that the fitted survival tree in the present study also found the age of 66 years as the cut-point in determining the risk factor of age and then spontaneously selected 79 years as the second cut-point; however, these cut-points have been recognized as important in the previous studies.[24]

According to tree growth after age, factors including immunity, history of high blood pressure, and cardiovascular problems were identified as the most important variables affecting patients' survival. It should be noted that all the variables identified in the tree were significant in regression and also important based on the random survival forest algorithm. The immunity variable was only involved in tree at older ages. It is necessary to point out that according to the procedure of the National Anti-Corona Headquarters in Iran, the priority of vaccination in the general population has been with older people. In other words, due to the low number of young immune patients, the immunity factor is split only in the old.

The results of this study also showed that immunity, regardless of the vaccine type, increases survival. This is evident in the higher Kaplan-Meier curve of immune patients. In addition, no significant difference was observed between the immunity caused by Sinopharm and Sputnik vaccines in the Iranian population. However, most of the immunized patients had been vaccinated with the Sinopharm vaccine, and based on the current data, the ideal condition for comparison is unavailable. On the other hand, a study in Bahrain found that in addition to the effectiveness of both vaccines in reducing the risk of death, people vaccinated with Sinopharm are at a higher risk of death. However, in the mentioned study, to compare vaccines, only age and sex variables were controlled in the model; while in the present study, in addition to more demographic variables, the history of underlying diseases has also been controlled.[28]

It should be noted that the efficacy of the COVID-19 vaccine is defined as the prevention of death, severe disease, or disease transmission, and any vaccine that can fulfill this need can be considered effective. Regarding efficacy against COVID-19 variants, it has been reported that vaccination was 85% effective against Alpha, 75% against Beta, 54% against Gamma, and 74% against Delta variant.[29] The present study also confirmed the effectiveness of Sinopharm. According to other studies reported by the World Health Organization (WHO), immunity from Sinopharm would prevent 78.1% of infection and 78.7% of hospitalization.[30, 31] Another study on hospitalized patients who were fully vaccinated with mRNA COVID-19 found that the odds of death and intubation were 67% lower.[32] Importantly, SARS-CoV-2 vaccines can effectively reduce deaths, severe cases, cases with symptoms, and the spread of SARS-CoV-2 infections worldwide. Therefore, the most important and urgent action in preventing the global pandemic is accelerating vaccination and improving vaccination coverage.

In the present study, the history of opium use was also identified to be associated with survival. Consistent with this study, there is general agreement on the effect of opium and long-term cigar smoking history on the conversion of COVID-19 disease into a severe form of the disease and higher mortality rates among drug users. Previous studies have shown that opium and its derivatives can reduce the immune system's strength by increasing cytokine secretion, especially interleukin. Long-term use of opium also affects the respiratory system, making addicts more susceptible to structural lung disease and dysfunction.[33] Although some previous studies have shown that the production of pro-inflammatory cytokines in opium addicts decreases, there are four reasons for the increase in mortality in COVID-19 infected opium addicts, including decreased IFNs expression, pulmonary edema, increased thrombotic factors, increased expression of Angiotensin-converting enzyme 2 (ACE2).[34]

In line with the present study, previous studies have shown that cancer patients with COVID-19 have a higher mortality rate.[35] This can be attributed to the suppression of the immune system, the concurrence of therapeutic conditions, and the dysfunction of the lungs, which is often present in lung malignancy. Patients with hematologic cancer and those undergoing chemotherapy were also at higher risk of death due to increased immune suppression.[36]

Consistent with the present study, other studies have found high blood pressure and cardiovascular disease, diabetes, chronic lung disease, kidney and liver disease to affect the covid-19 patients' survival and hospitalization.[37, 38] Therefore, people with these underlying diseases should be carefully monitored and prioritized in allocating preventive measures, including vaccination.

Gender, pregnancy, and blood disease history were removed from tree growth and the stepwise regression, having the least importance in the forest. This finding indicates the agreement of all three methods in estimating the variables' effect on survival. Although previous studies have looked at gender differences in incidence, hospitalization, and mortality, most studies have not examined the interaction between gender and smoking or drug use. In China, higher death rates in men have been attributed to a higher prevalence of smoking and alcohol consumption.[39] Further, the difference in the prevalence of chronic and underlying diseases between the two sexes should also be noted. In most previous studies, the incidence and mortality of COVID-19 in men have been reported more than in women due to the higher prevalence of related risk factors such as smoking, alcohol, and lack of hand hygiene, and the non-use of masks in men.[40] Therefore, more caution should be exercised when interpreting the role of gender in the incidence severe disease, hospitalization, and death from COVID-19 disease. The previous studies that examined the incidence, hospitalization, and death from COVID-19, from the perspective of biological gender differences, have attributed this difference to the predominant expression of the ACE2 enzyme in men.[41]

In addition, diabetes has been identified in most studies as a risk factor for COVID-19,[42, 43] whereas, in the present study, the effect of diabetes was insignificant in the multiple models. More examination of this variable in simple models that included only diabetes showed that this factor affects survival. Therefore, it can be concluded that the concurrence of diabetes and cardiovascular problems has led to a strong correlation between these two variables, confounding the relationship between diabetes and the survival of Covid-19 patients in multiple analyses.

In the Kaplan–Meier diagram of total patient survival, the chart's slope was very gentle after twenty-one days, indicating that after approximately three weeks, most patients experienced the outcome (including death/discharge). In a study in Italy on patients hospitalized from COVID-19, 22% of deaths occurred in the first 14 days after hospitalization and 27.6% after 30 days.[44] In other studies, median survival times of 13 and 20 days have been reported.[45] These differences can be due to the severity and type of virus variants, the condition of patients at the time of admission, having underlying diseases, and demographic factors, such as age and gender. In addition, the median length of hospital stay in the present study was estimated to be seven days. Other studies have reported different median and mean of hospital stays (in Ghana 13.8 days, in France 15.9 days,[46, 47] in China (Wuhan) 10-13 days, and in Vietnam 21 days.[48, 49] The characteristics of patients and different admission criteria in each country can be the reason for the different lengths of hospital stay. However, the short hospital stay in the present study compared to other areas can be attributed to the implementation of educational programs, active patient care, referral in the first days of the disease, and the implementation of telemedicine.[50]

Comparing the models, C-index depicted better performance for random survival forest. Furthermore, the variables identified in the random survival forest included those existing in the final Cox regression model. It is necessary to pay attention that the variables of little importance in the forest were not significant in the regression or did not stay in the final stepwise regression at all. The survival conditional tree branches were also a subset of variables identified by forest and Cox regression methods. Finally, there was general agreement among all methods in determining the effective variables, although the random survival forest identified a larger number of effective variables, which has been proven in other previous studies.[50]

Conclusion

The results of the present study showed that although age is the most effective factor in the survival of patients with COVID-19, immunity can control the hazard of this covariate. Notably, both Sinopharm and Sputnik vaccines reduced the hazard, and no significant difference was observed in their efficacy in real conditions. All models also identified individuals at higher risk with underlying diseases, including cancer, a history of chronic lung disease, kidney disease, high blood pressure, and cardiovascular disease. Therefore, it is recommended that prevention and treatment measures such as vaccination, disease detection, active care, hospitalization, and timely treatment of the older people and patients with any of the above underlying problems be prioritized.

Strong Points

The simultaneous application of three models of Cox regression, conditional tree, and random survival forest, which has made it possible to compare models, is the strong point of this study. So far, no studies have used survival tree and forest to estimate vaccination effectiveness.

Limitations of the Study

Lack of access to some clinical variables and missing data can be considered limitations of this study. Due to missing values, especially in length of hospital stay, patient's final status (death/discharge), and vaccination status, the data of 25% of hospitalized patients with low blood oxygen levels were not included in the analysis.

Acknowledgment

The authors of this study thank Fasa University of Medical Sciences for providing data. This study was approved and financially supported by the Research Vice-chancellor of Shiraz University of Medical Sciences (grant No. 22474). This study was approved by the Ethics Committee of Shiraz University of Medical Sciences (IR.SUMS.REC.1399.1251).

Ethics Approval and Consent to Participate

The authors have completely observed ethical issues, including plagiarism, informed consent, misconduct, data fabrication or falsification, double publication and submission, redundancy, and so on.

Availability of Supporting Data

Not available. The data of this study was extracted from the Covid-19 patient registration system in the hospital. Providing information requires permission from the authorities.

Authors' Contribution

Mozhgan Seif: final analysis, providing the main idea of study, and methodology. Mehdi Sharafi: data analysis, developing the idea, and revising the final manuscript. Marziyeh S. Safe and Mohsen Bayati: drafting the manuscript, Jalal Karimi: providing the data and contributing in drafing the manuscript. All authors approved the final version of the manuscript that was submitted.

Conflict of interest*:* None declared.

References

1 Liu W, Yang C, Liao YG, Wan F, Lin L, Huang

X, Zhang BH, et al. Risk factors for COVID-19 progression and mortality in hospitalized patients without pre-existing comorbidities. J Infect Public Health. 2022 Jan;15(1):13-20. doi: 10.1016/j.jiph.2021.11.012. Epub 2021 Nov 18. PMID: 34861603; PMCID: PMC8598256.

2 Moftakhar L, Seif M, Safe MS. Exponentially Increasing Trend of Infected Patients with COVID-19 in Iran: A Comparison of Neural Network and ARIMA Forecasting Models. Iran J Public Health. 2020 Oct;49(Suppl 1):92-100. doi: 10.18502/ijph.v49iS1.3675. PMID: 34268211; PMCID: PMC8266002.

3 Available from: https://www.who.int/publications/m/item/weekly-epidemiological-update-on-covid-19---18-january-2022.

4 Tsai PH, Lai WY, Lin YY, Luo YH, Lin YT, Chen HK, Chen YM, Lai YC, Kuo LC, Chen SD, Chang KJ, Liu CH, Chang SC, Wang FD, Yang YP. Clinical manifestation and disease progression in COVID-19 infection. J Chin Med Assoc. 2021 Jan 1;84(1):3-8. doi: 10.1097/JCMA.0000000000000463. PMID: 33230062.

5 Niroomand N, Bayati M, Seif M, Delavari S, Delavari S. Self-medication Pattern and Prevalence Among Iranian Medical Sciences Students. Curr Drug Saf. 2020;15(1):45-52. Doi: 10.2174/1574886314666191022095058. PMID: 31642791.

6 Ren X, Li Y, Yang X, Li Z, Cui J, Zhu A, Zhao H, Yu J, Nie T, Ren M, Dong S, Cheng Y, Chen Q, Chang Z, Sun J, Wang L, Feng L, Gao GF, Feng Z, Li Z. Evidence for pre-symptomatic transmission of coronavirus disease 2019 (COVID-19) in China. Influenza Other Respir Viruses. 2021 Jan;15(1):19-26. doi: 10.1111/irv.12787. Epub 2020 Aug 7. PMID: 32767657; PMCID: PMC7436222.

7 Cascella M, Rajnik M, Aleem A, Dulebohn SC, Di Napoli R. Features, Evaluation, and Treatment of Coronavirus (COVID-19). 2023 Jan 9. In: StatPearls [Internet]. Treasure Island (FL): StatPearls Publishing; 2023 Jan–. PMID: 32150360.

8 Ghaebi M, Tahmasebi S, Jozghorbani M, Sadeghi A, Thangavelu L, Zekiy AO, Esmaeilzadeh A. Risk factors for adverse outcomes of COVID-19 patients: Possible basis for diverse responses to the novel coronavirus SARS-CoV-2. Life Sci. 2021 Jul 15;277:119503. doi: 10.1016/j.lfs.2021.119503. Epub 2021 Apr 15. PMID: 33865882; PMCID: PMC8046708.

9 Gao YD, Ding M, Dong X, Zhang JJ, Kursat Azkur A, Azkur D, Gan H, Sun YL, Fu W, Li W, Liang HL, Cao YY, Yan Q, Cao C, Gao HY, Brüggen MC, et al. Risk factors for severe and critically ill COVID-19 patients: A review. Allergy. 2021 Feb;76(2):428-455. doi: 10.1111/all.14657. Epub 2020 Dec 4. PMID: 33185910.

10 Luo J, Zhu X, Jian J, Chen X, Yin K. Cardiovascular disease in patients with COVID-19: evidence from cardiovascular pathology to treatment. Acta Biochim Biophys Sin (Shanghai). 2021 Mar 2;53(3):273-282. doi: 10.1093/abbs/gmaa176. PMID: 33428706; PMCID: PMC7929476.

11 Muñoz-Rodríguez JR, Gómez-Romero FJ, Pérez-Ortiz JM, López-Juárez P, Santiago JL, Serrano-Oviedo L, et al; COVID-19 SESCAM Network. Characteristics and Risk Factors Associated With Mortality in a Multicenter Spanish Cohort of Patients With COVID-19 Pneumonia. Arch Bronconeumol. 2021 Apr;57:34-41. doi: 10.1016/j.arbres.2021.02.021. Epub 2021 Mar 9. PMID: 34629641; PMCID: PMC7939995.

12 Erdal GS, Polat O, Erdem GU, Korkusuz R, Hindilerden F, Yilmaz M, et al. The mortality rate of COVID-19 was high in cancer patients: a retrospective single-center study. Int J Clin Oncol. 2021 May;26(5):826-834. doi: 10.1007/s10147-021-01863-6. Epub 2021 Jan 24. PMID: 33486624; PMCID: PMC7826293.

13 Bowles KH, McDonald M, Barrón Y, Kennedy E, O'Connor M, Mikkelsen M. Surviving COVID-19 After Hospital Discharge: Symptom, Functional, and Adverse Outcomes of Home Health Recipients. Ann Intern Med. 2021 Mar;174(3):316-325. doi: 10.7326/M20-5206. Epub 2020 Nov 24. PMID: 33226861; PMCID: PMC7707212.

14 Ambale-Venkatesh B, Quinaglia T, Shabani M, Sesso J, Kapoor K, Matheson MB, et al. Prediction of Mortality in hospitalized COVID-19 patients in a statewide health network. medRxiv [Preprint]. 2021 Feb 19:2021.02.17.21251758. doi: 10.1101/2021.02.17.21251758. PMID: 33619510; PMCID: PMC7899480.

15 Ortiz RIH, Barrera JCB, Barrera KMB, editors. Analysis model of the most important factors in Covid-19 through data mining, descriptive statistics and random forest. 2020 IEEE International Autumn Meeting on Power, Electronics and Computing (ROPEC); 2020: IEEE. doi: 10.1109/ROPEC50909.2020.9258765.

16 Osi AA, Abdu M, Muhammad U, Ibrahim A, Isma'il LA, Suleiman AA, et al. A Classification Approach for Predicting COVID-19 Patient's Survival Outcome with Machine Learning Techniques. medRxiv. January 2022Science Forum (Journal of Pure and Applied Sciences) 22(1):1, doi: 10.5455/sf.aaosi2.

17 Alsulami M, Abu-Zinadah H, Ibrahim AH. Machine Learning Model and Statistical Methods for COVID-19 Evolution Prediction. Wireless Communications and Mobile Computing. 2021;2021. doi: 10.1155/2021/4840488.

18 Escobar-Agreda S, Silva-Valencia J, Rojas-Mezarina L, Vargas-Herrera J. Survival of health workers infected by SARS-CoV-2 in the context of vaccination against COVID-19 in Peru. medRxiv. 2021.doi: 10.1101/2021.08.03.21260614.

19 Scobie HM, Johnson AG, Suthar AB, Severson R, Alden NB, Balter S, et al. Monitoring Incidence of COVID-19 Cases, Hospitalizations, and Deaths, by Vaccination Status - 13 U.S. Jurisdictions, April 4-July 17, 2021. MMWR Morb Mortal Wkly Rep. 2021 Sep 17;70(37):1284-1290. doi: 10.15585/mmwr.mm7037e1. PMID: 34529637; PMCID: PMC8445374.

20 Štefan M, Dlouhý P, Bezdíčková L. [Vaccination

against COVID-19]. Klin Mikrobiol Infekc Lek. 2021 Jun;27(2):49-60. Czech. PMID: 34648642.

21 Muthukrishnan J, Vardhan V, Mangalesh S, Koley M, Shankar S, Yadav AK, et al. Vaccination status and COVID-19 related mortality: A hospital based cross sectional study. Med J Armed Forces India. 2021 Jul;77(Suppl 2):S278-S282. doi: 10.1016/j.mjafi.2021.06.034. Epub 2021 Jul 26. PMID: 34334894; PMCID: PMC8313045.

22 Dietz K, Gail M, Krickeberg K, Samet J, Tsiatis A. Statistics for biology and health. Survival Analysis, Edition Springer. 2002;7.

23 Yang Q, Li J, Zhang Z, Wu X, Liao T, Yu S, et al. Clinical characteristics and a decision tree model to predict death outcome in severe COVID-19 patients. BMC Infect Dis. 2021 Aug 9;21(1):783. doi: 10.1186/s12879-021-06478-w. PMID: 34372767; PMCID: PMC8351764.

24 An C, Lim H, Kim DW, Chang JH, Choi YJ, Kim SW. Machine learning prediction for mortality of patients diagnosed with COVID-19: a nationwide Korean cohort study. Sci Rep. 2020 Oct 30;10(1):18716. doi: 10.1038/s41598-020-75767-2. PMID: 33127965; PMCID: PMC7599238.

25 Mueller AL, McNamara MS, Sinclair DA. Why does COVID-19 disproportionately affect older people? Aging (Albany NY). 2020 May 29;12(10):9959-9981. doi: 10.18632/aging.103344. Epub 2020 May 29. PMID: 32470948; PMCID: PMC7288963.

26 Liu Y, Yan LM, Wan L, Xiang TX, Le A, Liu JM, Peiris M, Poon LLM, Zhang W. Viral dynamics in mild and severe cases of COVID-19. Lancet Infect Dis. 2020 Jun;20(6):656-657. doi: 10.1016/S1473-3099(20)30232-2. Epub 2020 Mar 19. PMID: 32199493; PMCID: PMC7158902.

27 Blagosklonny MV. From causes of aging to death from COVID-19. Aging (Albany NY). 2020 Jun 12;12(11):10004-10021. doi: 10.18632/aging.103493. Epub 2020 Jun 12. PMID: 32534452; PMCID: PMC7346074.

28 AlQahtani M, Bhattacharyya S, Alawadi A, Mahmeed HA, Sayed JA, Justman J, et al. Morbidity and mortality from COVID-19 post-vaccination breakthrough infections in association with vaccines and the emergence of variants in Bahrain. 2021. doi: 10.21203/rs.3.rs-828021/v1.

29 Hodgson SH, Mansatta K, Mallett G, Harris V, Emary KRW, Pollard AJ. What defines an efficacious COVID-19 vaccine? A review of the challenges assessing the clinical efficacy of vaccines against SARS-CoV-2. Lancet Infect Dis. 2021 Feb;21(2):e26-e35. doi: 10.1016/S1473-3099(20)30773-8. Epub 2020 Oct 27. PMID: 33125914; PMCID: PMC7837315.

30 WHO. Evidence Assessment: Sinopharm/BBIBP COVID-19 Vaccine. World Health Organization; 2021.

31 Shapiro J, Dean NE, Madewell ZJ, Yang Y, Halloran ME, Longini IM. Efficacy estimates for various COVID-19 vaccines: what we know from the literature and reports. medRxiv. 2021. doi: 10.1101/2021.05.20.21257461.

32 Tenforde MW, Self WH, Adams K, Gaglani M, Ginde AA, McNeal, et al. Influenza and Other Viruses in the Acutely Ill (IVY) Network. Association Between mRNA Vaccination and COVID-19 Hospitalization and Disease Severity. JAMA. 2021 Nov 23;326(20):2043-2054. doi: 10.1001/jama.2021.19499. PMID: 34734975; PMCID: PMC8569602.

33 Riahi T, Sadeghzadeh-Bazargan A, Shokri S, Ahmadvand D, Hassanlouei B, Baghestani A, Khazaeian A, Seifi Gharabaghloo F, Hassanzadeh M, Goodarzi A. The effect of opium on severity of COVID-19 infection:An original study from Iran. Med J Islam Repub Iran. 2021 Sep 8;35:115. doi: 10.47176/mjiri.35.115. PMID: 34956961; PMCID: PMC8683804.

34 Dolati-Somarin A, Abd-Nikfarjam B. The Reasons for Higher Mortality Rate in Opium Addicted Patients with COVID-19: A Narrative Review. Iran J Public Health. 2021 Mar;50(3):470-479. doi: 10.18502/ijph.v50i3.5587. PMID: 34178794; PMCID: PMC8214617.

35 Curigliano G. Cancer Patients and Risk of Mortality for COVID-19. Cancer Cell. 2020 Aug 10;38(2):161-163. doi: 10.1016/j.ccell.2020.07.006. Epub 2020 Jul 25. PMID: 32710820; PMCID: PMC7381394.

36 Addeo A, Friedlaender A. Cancer and COVID-19: Unmasking their ties. Cancer Treat Rev. 2020 Aug;88:102041. doi: 10.1016/j.ctrv.2020.102041. Epub 2020 Jun 3. PMID: 32516704; PMCID: PMC7831797.

37 Wu ZH, Tang Y, Cheng Q. Diabetes increases the mortality of patients with COVID-19: a meta-analysis. Acta Diabetol. 2021 Feb;58(2):139-144. doi: 10.1007/s00592-020-01546-0. Epub 2020 Jun 24. PMID: 32583078; PMCID: PMC7311595.

38 Li J, Huang DQ, Zou B, Yang H, Hui WZ, Rui F, et al. Epidemiology of COVID-19: A systematic review and meta-analysis of clinical characteristics, risk factors, and outcomes. J Med Virol. 2021 Mar;93(3):1449-1458. doi: 10.1002/jmv.26424. Epub 2020 Aug 25. PMID: 32790106; PMCID: PMC7436673.

39 Abate BB, Kassie AM, Kassaw MW, Aragie TG, Masresha SA. Sex difference in coronavirus disease (COVID-19): a systematic review and meta-analysis. BMJ Open. 2020 Oct 6;10(10):e040129. doi: 10.1136/bmjopen-2020-040129. PMID: 33028563; PMCID: PMC7539579.

40 Cai H. Sex difference and smoking predisposition in patients with COVID-19. Lancet Respir Med. 2020 Apr;8(4):e20. doi: 10.1016/S2213-2600(20)30117-X. Epub 2020 Mar 11. Erratum in: Lancet Respir Med. 2020 Apr;8(4):e26. PMID: 32171067; PMCID: PMC7103991.

41 Gebhard C, Regitz-Zagrosek V, Neuhauser HK, Morgan R, Klein SL. Impact of sex and gender on COVID-19 outcomes in Europe. Biol Sex Differ. 2020 May 25;11(1):29. doi: 10.1186/s13293-020-00304-9. PMID: 32450906; PMCID: PMC7247289.

42 Abdi A, Jalilian M, Sarbarzeh PA, Vlaisavljevic Z. Diabetes and COVID-19: A systematic review on the current evidences. Diabetes Res Clin Pract. 2020 Aug;166:108347. doi: 10.1016/j.diabres.2020.108347. Epub 2020 Jul 22. PMID: 32711003; PMCID: PMC7375314.

43 Peña JE, Rascón-Pacheco RA, Ascencio-Montiel IJ, González-Figueroa E, Fernández-Gárate JE, Medina-Gómez OS, Borja-Bustamante P, et al. Borja-Aburto VH. Hypertension, Diabetes and Obesity, Major Risk Factors for Death in Patients with COVID-19 in Mexico. Arch Med Res. 2021 May;52(4):443-449. doi: 10.1016/j.arcmed.2020.12.002. Epub 2020 Dec 16. PMID: 33380361; PMCID: PMC7832055.

44 Ferroni E, Giorgi Rossi P, Spila Alegiani S, Trifirò G, Pitter G, Leoni O, et al. Survival of Hospitalized COVID-19 Patients in Northern Italy: A Population-Based Cohort Study by the ITA-COVID-19 Network. Clin Epidemiol. 2020 Dec 8;12:1337-1346. doi: 10.2147/CLEP.S271763. PMID: 33335428; PMCID: PMC7737545.

45 Marimuthu Y, Kunnavil R, Anil NS, Nagaraja SB, Satyanarayana N, Kumar J, et al. Clinical profile and risk factors for mortality among COVID-19 inpatients at a tertiary care centre in Bengaluru, India. Monaldi Arch Chest Dis. 2021 May 17;91(3). doi: 10.4081/monaldi.2021.1724. PMID: 34006039.

46 Boëlle PY, Delory T, Maynadier X, Janssen C, Piarroux R, Pichenot M, et al. Trajectories of Hospitalization in COVID-19 Patients: An Observational Study in France. J Clin Med. 2020 Sep 29;9(10):3148. doi: 10.3390/jcm9103148. PMID: 33003375; PMCID: PMC7600846.

47 Ashinyo ME, Duti V, Dubik SD, Amegah KE, Kutsoati S, Oduro-Mensah E, et al. Clinical characteristics, treatment regimen and duration of hospitalization among COVID-19 patients in Ghana: a retrospective cohort study. Pan Afr Med J. 2020 Sep 15;37(Suppl 1):9. doi: 10.11604/pamj.supp.2020.37.9.25718. PMID: 33294110; PMCID: PMC7704352.

48 Thai PQ, Toan DTT, Son DT, Van HTH, Minh LN, Hung LX, te al. Factors associated with the duration of hospitalisation among COVID-19 patients in Vietnam: A survival analysis. Epidemiol Infect. 2020 Jun 10;148:e114. doi: 10.1017/S0950268820001259. PMID: 32517822; PMCID: PMC7306545.

49 Shao F, Xu S, Ma X, Xu Z, Lyu J, Ng M, te al. In-hospital cardiac arrest outcomes among patients with COVID-19 pneumonia in Wuhan, China. Resuscitation. 2020 Jun;151:18-23. doi: 10.1016/j.resuscitation.2020.04.005. Epub 2020 Apr 10. PMID: 32283117; PMCID: PMC7151543.

50 Seif M, Sharafi M, Ghaem H, Kasraei F. Factors associated with survival of Iranian patients with COVID-19: comparison of Cox regression and mixture cure model. Trop Dis Travel Med Vaccines. 2022 Mar 1;8(1):4. doi: 10.1186/s40794-022-00162-w. PMID: 35227332; PMCID: PMC8885138.