Archives of Anesthesiology and Critical Care (Autumn 2023); 9(Supplement 2): 479-487.

Available online at http://aacc.tums.ac.ir

Research Article

# Predicting Mortality and ICUs Transfer in Hospitalized COVID-19 Patients Using Random Forest Model

**Roya Najafi-Vosough[1], Mohammad Hossein Bakhshaei[2]*, Mahnaz Farzian[3]**

[1]Department of Biostatistics, School of Public Health, Hamadan University of Medical Sciences, Hamadan, Iran.

[2]Department of Anaesthesiology, School of Medicine, Hamadan University of Medical Sciences, Hamadan, Iran.

[3]Department of Nursing, Hamadan University of Medical Sciences, Hamadan, Iran.

**ARTICLE INFO**

**Article history:**
Received 19 November 2022
Revised 11 December 2022
Accepted 25 December 2022

**Keywords:**
COVID-19;
Mortality;
Intensive care units;
Random forest;
Logistic regression

**ABSTRACT**

**Background:** The objective of the present study was to identify prognostic factors associated with mortality and transfer to intensive care units (ICUs) in hospitalized COVID-19 patients using random forest (RF). Also, its performance was compared with logistic regression (LR).

**Methods:** In this retrospective cohort study, information of 329 COVID-19 patients were analyzed. These patients were hospitalized in Besat hospital in Hamadan province, the west of Iran. The RF and LR models were used for predicting mortality and transfer to ICUs. These models' performance was assessed using area under the receiver operating characteristic curve (AUC) and accuracy.

**Results:** Of the 329 COVID-19 patients, 57 (15.5%) patients died and 106 (32.2%) patients were transferred to ICUs. Based on multiple LR model, there was a significant association between age (OR=1.02; 95% CI=1.00-1.05), cough (OR=0.24; 95% CI=0.10-0.56), and ICUs (OR=7.20; 95% CI=3.30-15.69) with death. Also, a significant association was found between kidney disease (OR=3.90; 95% CI=1.04-14.63), decreased sense of smell (OR=0.28; 95% CI=0.10-0.73), Kaletra (OR=2.53; 95% CI=1.39-4.59), and intubation (OR=8.32; 95% CI=3.80-18.24) with transfer to ICUs. RF showed that the order of variable importance has belonged to age, ICUs, and cough for predicting mortality; and age, intubation, and Kaletra for predicting transfer to ICUs.

**Conclusion:** This study showed that the performance of RF provided better results compared to LR for predicting mortality and ICUs transfer in hospitalized COVID-19 patients.

## Introduction

Here In late 2019, a new type of Coronavirus (COVID-19) called Severe Acute Respiratory Syndrome Coronavirus 2 (SARS-CoV-2) first appeared in Wuhan, China. Since then, it has rapidly spread around the world and become a worldwide epidemic [1-2]. The SARS-CoV-2 infection has led to extensive morbidity and mortality throughout the world [3]. Approximately 20% to 30% of patients experience a moderate to severe form of the disease that may lead to hospitalization or death [4]. In addition, 5% to 12% of all COVID-19 patients and up to 33% of hospitalized patients require intensive care units (ICUs) [4-6].

Studies have shown that most COVID-19 patients experience mild symptoms such as fever, dry cough, weakness, shortness of breath, and diarrhea [7]. However, elderly patients and those with diabetes, hypertension, and cardiovascular disease may experience worse outcomes, such as hospitalization [8-9]. Some features, such as age, body temperature, lymphocyte

The authors declare no conflicts of interest.

*Corresponding author.
E-mail address: bakhshaei@umsha.ac.ir

count, and lung imaging features, have been reported as the most common predictors in the diagnosis and prognosis of COVID-19 [10]. Several prognostic factors such as older age, male sex, obesity, smoking, cardiovascular disease, diabetes, hypertension, and chronic respiratory disease are related to severe illness and death from COVID-19. However, the results of several studies have been inconsistent [11-15].

In the past decade, machine learning models have been widely applied to many fields [16-20]. Among them, the random forests (RF) is well-known as one of the best models, which is widely employed for classification and regression problems. This model can improve predictions by considering the complex nonlinear relationships between variables [19-21]. However, RF cannot determine the direction of association of variables on outcome. To overcome this problem can be applied to the logistic regression (LR) model [22].

Despite the several studies that have used machine learning models for predicting the prognosis of COVID-19 patients [10, 14, 15, 23-28], only a few of them have been performed in Iran [29-31]. Furthermore, COVID-19 is still widely spreading in Asian countries such as Iran. Hence, this study's main goal was to identify prognostic factors associated with mortality and transfer to intensive care units (ICUs) in hospitalized COVID-19 patients using RF. Also, its performance was compared with LR.

## Methods

In this retrospective cohort study, information of 329 COVID-19 patients was analyzed. These patients were hospitalized in Besat hospital in Hamadan, Iran. Data were extracted from hospital records using a checklist of items according to the context of the patient's records. The checklist included data on demographic variables, medical history, co-morbidities, laboratory tests, and clinical symptoms. Missing values of these variables were imputed with the multiple imputations method. The outcome was mortality and transfer to ICUs.

RF is a tree-based model that was applied to determine important prognostic factors. This model has proved its success in classification and regression problems. In RF, data were randomly divided into two parts. The learning part (2/3 of the data set) was used for developing the model, and the test part was used (1/3 of the data set) to check the data validity. Totally, 1000 bootstrap samples were constructed from the learning part. Then decision trees for each bootstrap sample were grown. In each tree node, a subset of p variables was selected randomly and among which, the best variables are chosen for splitting. The RF Final estimate was the average of all results from each tree. The variables were ranked using the Gini Index [21]. LR model was performed to assess the effect of prognostic factors on mortality and transfer to ICUs. In the univariate LR model, significant variables were selected and entered into the multiple LR model. Crude and adjusted odds ratios (ORs) were reported to address the effect of prognosis factors associated with mortality and transfer to ICUs. A p-value less than 0.05 was considered statistically significant.

The cross-validation method was used in the performance evaluation of RF and LR models, in which the dataset was randomly divided into training (70%) and test (30%) sets. Then, the discrimination ability of both models was assessed using the area under the receiver operating characteristic curve (AUC) and accuracy. This procedure was repeated 100 times and the average values of AUC and accuracy were computed.

The statistical analyses were performed using R Version 3.6.3 [32], with the following packages "randomForest", "CORElearn" and " pROC".

## Results

During the study period, 57 (15.5%) patients died and 106 (32.2%) patients were transferred to ICUs. The mean age (standard deviation) of all patients was 43.11(28.89) years old. The majority of them were men (60.9%) and married (65.5%). The characteristics of COVID-19 patients and results from the univariate LR analysis of the prognostic factors associated with mortality and transfer to ICUs are given in (Table 1).

Data are expressed as Mean (SD) or N (%); CI: confidence interval.

The results of the multiple LR model for mortality and transfer to ICUs are presented in (Table 2-3), respectively.

The risk of death increased significantly with age (P<0.001). There was a significant association between marital and death, and the OR of death in COVID-19 married patients was 2.10 (95% CI: 1.03-4.28) compared to single patients. The OR estimate of death among patients with heart disease was 3.14 (95% CI=1.37-7.19) compared to those patients without heart disease. The COVID-19 patients who had symptoms such as fever (P=0.001), headache (P=0.017), and cough (P<0.001) were significantly at a lower risk of death. However, the patients who had a weakness (P<0.001) were significantly at a higher risk of death. Compared to COVID-19 patients without hypertension, the OR estimate of death was 2.40 (95% CI=1.31-4.40) in patients with hypertension. There was an association between WBC and death. The OR estimate of the death of COVID-19 patients was 1.98 (95% CI=1.08-3.63) for WBC >=11000 compared to WBC <11000 (Table 1). Transfer to ICUs was the strongest risk factor for COVID-19 patients' mortality. The OR estimate of death in those who transfer to ICUs was 5.72 (95% CI=3.01-0.86) compared to those without transfer to ICUs. After controlling for the potential confounding effect of other variables, the association became stronger and reached 7.87 (95% CI=3.47-17.86).

There was an association between age and transfer to ICUs in COVID-19 patients (p=0.032). Compared to married COVID-19 patients, single patients were at a lower risk of transfer to ICUs(p=0.009). There was a

direct association between transfer to ICUs and WBC. Compared to COVID-19 patients with WBC <11000, the OR estimate of transfer to ICUs was 2.14 (95% CI=1.32-3.47) in patients with WBC >=11000. The COVID-19 patients who had headaches (P=0.008), and decreased sense of smell (P<0.001) were significantly at a lower risk of transfer to ICUs (Table 1). The risk of transfer to ICUs was higher in COVID-19 patients with kidney disease than in those without kidney disease (p=0.043). Treatment with Kaletra increased the risk of transfer to ICUs in COVID-19 patients (P=0.002). This OR estimate increased to 2.53 (95% CI=1.39-4.59) when adjusted for other variables. Intubation was the strongest risk factor for transfer to ICUs in COVID-19 patients. The OR estimate of transfer to ICUs in those who had intubation was 5.59 (95% CI=2.90-10.78) compared to those without intubation. After controlling for the potential confounding effect of other variables, the association became stronger and reached 8.32 (95% CI=3.80-18.24). (Figure 1) displays the top ten variable importance obtained from RF for mortality and transfer to ICUs in COVID-19 patients.

The results showed that age, ICUs, and cough as the three most important variables for predicting mortality in COVID-19 patients (Figure-1 (A)). Also, RF identified that age, intubation, and Kaletra as the three most important variables for predicting transfer to ICUs in COVID-19 patients (Figure 1 (B)).

(Table 4) shows the performance of LR and RF models for Predicting mortality and ICUs transfer in hospitalized COVID-19 patients in training and testing datasets. As seen, the performance of RF compared to LR was better for both outcomes in the training and testing datasets, respectively.

**Table 1- The characteristics of COVID-19 patients and results of univariate logistic regression model for mortality and transfer to ICUs.**

| Variables | Patient status | | | | ICUs Transfer | | | |
|---|---|---|---|---|---|---|---|---|
| | **Died N=51** | **Alive N=278** | **Unadjusted OR (95% CI)** | **P value** | **Yes N=106** | **No N=223** | **Unadjusted OR (95% CI)** | **P value** |
| Age (Year) | 56.67 (24.74) | 40.62 (28.94) | 1.02 (1.00-1.03) | <0.001 | 38.14 (30.77) | 45.47 (27.71) | 0.99 (0.98-0.99) | 0.032 |
| Sex | | | | | | | | |
| Male | 36 (70.6) | 164 (59.0) | 1 | | 63 (60.6) | 135 (61.1) | 1 | |
| Female | 15 (29.4) | 114 (41.0) | 0.59 (0.31-1.14) | 0.122 | 41 (49.4) | 86 (38.9) | 1.08 (0.67-1.74) | 0.728 |
| Marital status | | | | | | | | |
| Single | 11 (21.6) | 102 (36.7) | 1 | | 47 (46.1) | 63 (29.0) | 1 | |
| Married | 40 (78.4) | 176 (63.3) | 2.10 (1.03-4.28) | 0.040 | 55 (53.9) | 154 (71.0) | 0.52 (0.32-0.85) | 0.009 |
| Heart disease | | | | | | | | |
| No | 41 (80.4) | 258 (92.8) | 1 | | 94 (88.7) | 205 (91.9) | 1 | |
| Yes | 10 (19.6) | 20 (7.2) | 3.14 (1.37-7.19) | 0.007 | 12 (11.3) | 18 (8.1) | 1.45 (0.63-3.14) | 0.341 |
| Hypertension | | | | | | | | |
| No | 25 (49.0) | 194 (69.8) | 1 | | 72 (67.9) | 147 (65.9) | 1 | |
| Yes | 26 (51.0) | 84 (30.3) | 2.40 (1.31-4.40) | 0.005 | 34 (32.1) | 76 (34.1) | 0.91 (0.55-1.49) | 0.719 |
| Respiratory disease | | | | | | | | |
| No | 48 (94.1) | 265 (95.3) | 1 | | 102 (96.2) | 211 (94.6) | 1 | |
| Yes | 3 (5.9) | 13 (4.7) | 1.27 (0.35-4.64) | 0.713 | 4 (3.8) | 12 (5.4) | 0.69 (0.21-2.19) | 0.529 |
| Kidney disease | | | | | | | | |
| No | 48 (94.1) | 268 (96.4) | 1 | | 98 (92.5) | 218 (97.8) | 1 | |
| Yes | 3 (5.9) | 10 (3.6) | 1.67 (0.44-6.31) | 0.446 | 8 (7.5) | 5 (2.2) | 3.55 (1.13-1.15) | 0.029 |
| Liver disease | | | | | | | | |
| No | 49 (96.1) | 272 (97.8) | 1 | | 104 (98.1) | 217 (97.3) | 1 | |
| Yes | 2 | 6 | 1.85 | 0.459 | 2 | 6 | 0.69 | 0.660 |

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | (3.9) | (2.2) | (0.36-9.43) | | (1.9) | (2.7) | (0.13-3.50) | |
| Digestive disease | | | | | | | | |
| No | 50 | 242 | 1 | | 99 | 193 | 1 | |
| | (98.0) | (87.1) | | | (93.4) | (86.5) | | |
| Yes | 1 | 36 | 0.13 | 0.050 | 7 | 30 | 0.45 | 0.072 |
| | (2.0) | (12.9) | (0.01-1.00) | | (6.6) | (13.5) | (0.19-1.07) | |
| Diabetes | | | | | | | | |
| No | 39 | 231 | 1 | | 87 | 183 | 1 | |
| | (76.5) | (83.1) | | | (82.1) | (82.1) | | |
| Yes | 12 | 47 | 1.51 | 0.260 | 19 | 40 | 0.99 | 0.998 |
| | (23.5) | (16.9) | (0.73-3.10) | | (17.9) | (17.9) | (0.54-1.82) | |
| Anemia | | | | | | | | |
| No | 48 | 267 | 1 | | 99 | 212 | 1 | |
| | (94.1) | (96.0) | | | (95.2) | (95.9) | | |
| Yes | 3 | 11 | 1.51 | 0.534 | 5 | 9 | 1.17 | 0.775 |
| | (5.9) | (4.0) | (0.40-5.64) | | (4.8) | (4.1) | (0.38-3.60) | |
| LDH | | | | | | | | |
| <400 | 11 | 55 | 1 | | 6 | 20 | 1 | |
| | (21.6) | (19.8) | | | (12.2) | (21.1) | | |
| >=400 | 40 | 223 | 0.89 | 0.770 | 43 | 75 | 1.22 | 0.505 |
| | (78.4) | (80.2) | (0.43-1.86) | | (87.8) | (78.9) | (0.67-2.20) | |
| WBC | | | | | | | | |
| <11000 | 27 | 192 | 1 | | 58 | 159 | 1 | |
| | (52.9) | (69.1) | | | (55.2) | (73.6) | | |
| >=11000 | 24 | 86 | 1.98 | 0.027 | 47 | 57 | 2.14 | 0.002 |
| | (47.1) | (30.9) | (1.08-3.63) | | (44.8) | (26.4) | (1.32-3.47) | |
| ESR | | | | | | | | |
| <50 | 33 | 202 | 1 | | 51 | 88 | 1 | |
| | (64.7) | (72.7) | | | (68.0) | (70.4) | | |
| >=50 | 18 | 76 | 1.45 | 0.249 | 24 | 37 | 1.12 | 0.721 |
| | (35.3) | (27.3) | (0.77-2.72) | | (32.0) | (29.6) | (0.67-1.86) | |
| Blood pressure | | | | | | | | |
| Normal | 49 | 272 | 1 | | 94 | 212 | 1 | |
| | (96.1) | (97.8) | | | (98.9) | (97.2) | | |
| Abnormal | 2 | 6 | 1.85 | 0.459 | 1 | 6 | 0.69 | 0.660 |
| | (3.9) | (2.2) | (0.36-9.43) | | (1.1) | (2.8) | (0.13-3.50) | |
| PR | | | | | | | | |
| Normal | 47 | 254 | 1 | | 94 | 206 | 1 | |
| | (92.2) | (91.3) | | | (89.5) | (92.4) | | |
| Abnormal | 4 | 24 | 0.90 | 0.853 | 11 | 17 | 1.40 | 0.405 |
| | (7.8) | (8.7) | (0.29-2.71) | | (10.5) | (7.6) | (0.63-3.11) | |
| SPO2 | | | | | | | | |
| Normal | 47 | 268 | 1 | | 98 | 216 | 1 | |
| | (92.2) | (96.4) | | | (93.3) | (96.9) | | |
| Abnormal | 4 | 10 | 2.28 | 0.178 | 7 | 7 | 2.18 | 0.155 |
| | (7.8) | (3.6) | (0.68-7.57) | | (6.7) | (3.1) | (0.74-6.38) | |
| Fever | | | | | | | | |
| No | 30 | 95 | 1 | | 48 | 77 | 1 | |
| | (58.8) | (34.2) | | | (45.3) | (34.5) | | |
| Yes | 21 | 183 | 0.36 | 0.001 | 58 | 146 | 0.63(0.39- | 0.061 |
| | (41.2) | (65.8) | (0.19-0.66) | | (54.7) | (65.5) | 1.02) | |
| Headache | | | | | | | | |
| No | 40 | 168 | 1 | | 78 | 130 | 1 | |
| | (78.4) | (60.4) | | | (73.6) | (58.3) | | |
| Yes | 11 | 110 | 0.42 | 0.017 | 28 | 93 | 0.50 | 0.008 |
| | (21.6) | (39.6) | (0.20-0.85) | | (26.4) | (41.7) | (0.30-0.83) | |
| Sore throat | | | | | | | | |
| No | 45 | 234 | 1 | | 94 | 185 | 1 | |
| | (88.2) | (84.2) | | | (88.7) | (83.0) | | |

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| Yes | 6 (11.8) | 44 (15.8) | 0.70 (0.28-1.76) | 0.459 | 12 (11.3) | 38 (17.0) | 0.62 (0.31-1.24) | 0.180 |
| Decreased sense of smell | | | | | | | | |
| No | 42 (82.4) | 227 (81.7) | 1 | | 99 (93.4) | 170 (76.2) | 1 | |
| Yes | 9 (17.6) | 51 (18.3) | 0.95 (0.43-2.08) | 0.906 | 7 (6.6) | 53 (23.8) | 0.22 (0.09-0.51) | <0.001 |
| Gastrointestinal disorders | | | | | | | | |
| No | 44 (86.3) | 207 (74.5) | 1 | | 85 (80.2) | 166 (74.4) | 1 | |
| Yes | 7 (13.7) | 71 (25.5) | 0.46 (0.20-1.07) | 0.074 | 21 (19.8) | 57 (25.6) | 0.72 (0.40-1.26) | 0.253 |
| Shortness of breath | | | | | | | | |
| No | 24 (47.1) | 108 (38.8) | 1 | | 43 (40.6) | 89 (39.9) | 1 | |
| Yes | 27 (52.9) | 170 (61.2) | 0.71 (0.39-1.30) | 0.273 | 63 (59.4) | 134 (60.1) | 0.97 (0.60-1.55) | 0.910 |
| Cough | | | | | | | | |
| No | 39 (76.5) | 118 (42.4) | 1 | | 51 (48.1) | 106 (47.5) | 1 | |
| Yes | 12 (23.5) | 160 (57.6) | 0.27 (0.11-0.45) | <0.001 | 55 (51.9) | 117 (52.5) | 0.97 (0.61-1.55) | 0.922 |
| Diarrhea | | | | | | | | |
| No | 49 (96.1) | 237 (85.3) | 1 | | 95 (89.6) | 191 (85.7) | 1 | |
| Yes | 2 (3.9) | 41 (14.7) | 0.23 (0.05-1.00) | 0.051 | 11 (10.4) | 32 (14.3) | 0.69 (0.33-1.43) | 0.320 |
| Weakness | | | | | | | | |
| No | 35 (68.6) | 249 (89.6) | 1 | | 88 (83.0) | 196 (87.9) | 1 | |
| Yes | 16 (31.4) | 29 (10.4) | 3.92 (1.93-7.94) | <0.001 | 18 (17.0) | 27 (12.1) | 1.48 (0.77-2.83) | 0.231 |
| Myalgia | | | | | | | | |
| No | 50 (98.0) | 248 (89.2) | 1 | | 98 (92.5) | 200 (89.7) | 1 | |
| Yes | 1 (2.0) | 30 (10.8) | 0.16 (0.02-1.24) | 0.080 | 8 (7.5) | 23 (10.3) | 0.71 (0.30-1.64) | 0.424 |
| Hydroxychloroquine | | | | | | | | |
| No | 24 (47.1) | 94 (33.8) | 1 | | 34 (32.1) | 84 (37.3) | 1 | |
| Yes | 27 (52.9) | 184 (66.2) | 0.57 (0.31-1.05) | 0.072 | 72 (67.9) | 139 (62.3) | 1.28 (0.78-2.08) | 0.323 |
| Azithromycin | | | | | | | | |
| No | 36 (70.6) | 234 (84.2) | 1 | | 83 (78.3) | 187 (83.9) | 1 | |
| Yes | 15 (29.4) | 44 (15.8) | 2.21 (1.11-4.38) | 0.022 | 23 (21.7) | 36 (16.1) | 1.43 (0.80-2.58) | 0.221 |
| Kaletra | | | | | | | | |
| No | 36 (70.6) | 206 (74.1) | 1 | | 66 (62.3) | 176 (78.9) | 1 | |
| Yes | 15 (29.4) | 72 (25.9) | 1.19 (0.61-2.30) | 0.601 | 40 (37.3) | 47 (21.1) | 2.27 (1.36-3.77) | 0.002 |
| Intubation | | | | | | | | |
| No | 3 (5.9) | 278 (100) | - | - | 74 (69.8) | 207 (92.8) | 1 | |
| Yes | 48 (94.1) | 0 | - | - | 32 (30.2) | 16 (7.2) | 5.59 (2.90-10.78) | <0.001 |

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| ICUs | | | | | | | | |
| No | 17 (33.3) | 206 (74.1) | 1 | | - | - | - | - |
| Yes | 34 (66.7) | 72 (25.9) | 5.72 (3.01-0.86) | <0.001 | - | - | - | - |

**Table 2- The odds ratio (OR) estimates of Covid-19 patients mortality by different variable using multiple logistic regression model.**

| Variables | Adjusted OR(95% CI) | P value |
|---|---|---|
| Age (Year) | 1.02 (1.00-1.05) | 0.044 |
| Marital status | | |
| Single | 1 | |
| Married | 0.70 (0.17-2.83) | 0.626 |
| Heart disease | | |
| No | 1 | |
| Yes | 1.04 (0.33-3.21) | 0.940 |
| Hypertension | | |
| No | 1 | |
| Yes | 1.71 (0.66-4.43) | 0.267 |
| WBC | | |
| <11000 | 1 | |
| >=11000 | 1.91 (0.89-4.08) | 0.095 |
| Fever | | |
| No | 1 | |
| Yes | 0.76 (0.33-1.75) | 0.531 |
| Headache | | |
| No | 1 | |
| Yes | 0.71 (0.29-1.76) | 0.467 |
| Cough | | |
| No | 1 | |
| Yes | 0.24 (0.10-0.56) | 0.001 |
| Weakness | | |
| No | 1 | |
| Yes | 1.58 (0.61-4.11) | 0.342 |
| Azithromycin | | |
| No | 1 | |
| Yes | 1.37 (0.56-3.37) | 0.483 |
| ICUs | | |
| No | 1 | |
| Yes | 7.20 (3.30-15.69) | <0.001 |

CI: confidence interval.

**Table 3- The odds ratio (OR) estimates of Covid-19 patients transfer to ICUs by different variable using multiple logistic regression model.**

| Variables | Adjusted OR (95% CI) | P value |
|---|---|---|
| Age (Year) | 0.99 (0.97-1.00) | 0.220 |
| Marital status | | |
| Single | 1 | |
| Married | 0.69 (0.27-1.76) | 0.447 |
| Kidney disease | | |
| No | 1 | |
| Yes | 3.90 (1.04-14.63) | 0.043 |
| WBC | | |
| <11000 | 1 | |
| >=11000 | 1.62 (0.94-2.81) | 0.080 |
| Headache | | |
| No | 1 | |
| Yes | 0.93 (0.51-1.72) | 0.840 |
| Decreased sense of smell | | |

| | | |
|---|---|---|
| No | 1 | |
| Yes | 0.28 (0.10-0.73) | 0.009 |
| Kaletra | | |
| No | 1 | |
| Yes | 2.53 (1.39-4.59) | 0.002 |
| Intubation | | |
| No | 1 | |
| Yes | 8.32 (3.80-18.24) | <0.001 |

CI: confidence interval.

**Table 4- The performance criteria of models**

| | | Mortality | | Transfer to ICUs | |
|---|---|---|---|---|---|
| **Models** | **Data set** | **AUC** | **Accuracy** | **AUC** | **Accuracy** |
| Logistic regression | Train | 0.93±0.01 | 0.90±0.01 | 0.85±0.03 | 0.83±0.02 |
| | Test | 0.76±0.06 | 0.83±0.03 | 0.83±0.07 | 0.81±0.04 |
| Random Forest | Train | 0.97±0.01 | 0.87±0.01 | 0.88±0.04 | 0.88±0.02 |
| | Test | 0.82±0.04 | 0.85±0.03 | 0.87±0.05 | 0.89±0.04 |

AUC: Area under the receiver operating characteristic curve.

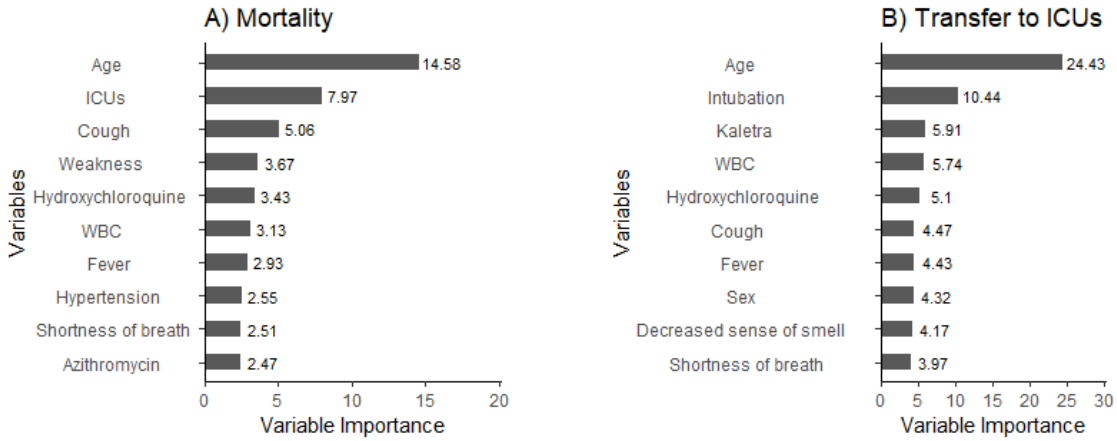


**Figure 1- The top ten variable importance for mortality and ICUs transfer in hospitalized COVID-19 patients.**

## *Discussion*

In this study, the effects of several variables on mortality and transfer to ICUs were assessed using RF and LR models. The results of RF indicated that age, ICUs, and cough as the top three most important variables for predicting death; and age, intubation, and Kaletra for predicting transfer to ICUs. Also, the multiple LR model indicated a significant association between age, cough, and ICUs with death; and kidney disease, decreased sense of smell, Kaletra, and intubation with transfer to ICUs.

We found a strong association between age and death. So that the risk of death increased with age. Also, the results of RF showed that age was an important variable for death and transfer to ICUs in hospitalized COVID-19 patients. These results are in agreement with previous studies [33-35]. Zhao et al. also reported that age was an important risk factor for death and transfer to ICUs in hospitalized COVID-19 patients [33].

According to the results, kidney disease was another important risk factor for transfer to ICUs. in COVID-19 patients. Subudhi et al. and Flythe et al. reported that kidney disease can lead to a severe prognosis of COVID-19 and even death in them [3, 36]. In this study, intubation was another significant variable for predicting transfer to ICUs in COVID-19 patients. This may be due to disease progression and worsening symptoms in these patients. Previous studies have confirmed this result in these patients [37-38].

In this study, we also compared the performance of RF and LR models using cross-validation. According to the results, the comparison between these models showed that the performance of RF was better than LR for predicting mortality and ICUs transfer in hospitalized COVID-19 patients. Various studies have investigated mortality and ICUs transfer among COVID-19 patients using different models. Yadaw et al. [39] compared the performance of LR, RF, support vector machine, and XGBoost to predict COVID-19 mortality. They found that XGBoost had better performance than other models. In another study, Gao et al. [40] showed that the neural network had the highest performance for predicting

COVID-19 patient's physiological deterioration and death. Moulaei et al. [38] also found the performance of RF was better than LR for predicting COVID-19 mortality. This finding is consistent with our results.

Based on our analysis, the RF had better performance than LR in predicting the ICUs transfer in COVID-19 patients. These results are in agreement with previous studies [24, 41-42]. Subudhi et al. compared the performance of machine learning algorithms for predicting ICU admission and mortality in COVID-19 patients. They found that ensemble-based methods had better performance than other machine learning algorithms [3]. Similar results were also reported in a study conducted by Orooji et al. [31].

The limitations of this study were the retrospective design and the absence of some data in patients' records. Despite these limitations, our study showed that RF had a good performance in predicting mortality and transfer to ICUs in COVID-19 patients. The results of this study can help better manage patients with COVID-19.

## Conclusion

This study showed that the performance of RF provided better results compared to LR for predicting mortality and ICUs transfer in hospitalized COVID-19 patients.

### Acknowledgements

The authors would like to express gratitude to the Vice-Chancellor of Research and Technology, Hamadan University of Medical Sciences for the approval and support of this study (Ethical code: IR.UMSHA.REC.1399.118). It is also necessary to thank the staff of Besat hospital for their cooperation.

### References

[1] Organization WH. Coronavirus disease 2019 (COVID-19): situation report, 73 2020 [Available from: Availableonline:https://www.who.int/emergencies/diseases/novel-coronavirus-2019/situation-reports.

[2] Yeşilkanat CM. Spatio-temporal estimation of the daily cases of COVID-19 in worldwide using random forest machine learning algorithm. Chaos Solitons Fractals. 2020; 140:110210.

[3] Subudhi S, Verma A, Patel AB, Hardin CC, Khandekar MJ, Lee H, et al. Comparing machine learning algorithms for predicting ICU admission and mortality in COVID-19. NPJ digital medicine. 2021;4(1):1-7.

[4] Covid C, Team R, COVID C, Team R, COVID C, Team R, et al. Severe outcomes among patients with coronavirus disease 2019 (COVID-19)—United States, February 12–March 16, 2020. MMWR Morb Mortal Wkly Rep. 2020; 69(12):343-6.

[5] Myers LC, Parodi SM, Escobar GJ, Liu VX. Characteristics of hospitalized adults with COVID-19 in an integrated health care system in California. Jama. 2020; 323(21):2195-8.

[6] Richardson S, Hirsch JS, Narasimhan M, Crawford JM, McGinn T, Davidson KW, et al. Presenting characteristics, comorbidities, and outcomes among 5700 patients hospitalized with COVID-19 in the New York City area. Jama. 2020; 323(20):2052-9.

[7] Hui DS, Azhar EI, Madani TA, Ntoumi F, Kock R, Dar O, et al. The continuing 2019-nCoV epidemic threat of novel coronaviruses to global health—The latest 2019 novel coronavirus outbreak in Wuhan, China. Int J Infect Dis. 2020; 91:264-6.

[8] Huang C, Wang Y, Li X, Ren L, Zhao J, Hu Y, et al. Clinical features of patients infected with 2019 novel coronavirus in Wuhan, China. Lancet. 2020; 395(10223):497-506.

[9] Chen N, Zhou M, Dong X, Qu J, Gong F, Han Y, et al. Epidemiological and clinical characteristics of 99 cases of 2019 novel coronavirus pneumonia in Wuhan, China: a descriptive study. Lancet. 2020; 395(10223):507-13.

[10] Chen Y, Liu Z, Li X, Zhao J, Wu D, Xiao M, et al. Risk factors for mortality due to COVID-19 in intensive care units: a single-center study. Ann Transl Med. 2021; 9(4):276.

[11] Du Y, Tu L, Zhu P, Mu M, Wang R, Yang P, et al. Clinical features of 85 fatal cases of COVID-19 from Wuhan. A retrospective observational study. Am J Respir Crit Care Med. 2020 Jun 1;201(11):1372-9.

[12] Liang W-h, Guan W-j, Li C-c, Li Y-m, Liang H-r, Zhao Y, et al. Clinical characteristics and outcomes of hospitalised patients with COVID-19 treated in Hubei (epicentre) and outside Hubei (non-epicentre): a nationwide analysis of China. Eur Respir J. 2020; 55(6):2000562.

[13] García LF, Gutiérrez ABP, Bascones MG. Relationship between obesity, diabetes and ICU admission in COVID-19 patients. Med Clin (English Ed). 2020;155(7):314.

[14] Jackson BR, Gold JA, Natarajan P, Rossow J, Neblett Fanfair R, da Silva J, et al. Predictors at admission of mechanical ventilation and death in an observational cohort of adults hospitalized with COVID-19. Clin infect dis. 2021. 73(11):e4141-e4151.

[15] Venturini S, Orso D, Cugini F, Crapis M, Fossati S, Callegari A, et al. Classification and analysis of outcome predictors in non-critically ill COVID-19 patients. Intern Med J. 2021; 51(4):506-14.

[16] Jeung M, Baek S, Beom J, Cho KH, Her Y, Yoon K. Evaluation of random forest and regression tree methods for estimation of mass first flush ratio in urban catchments. Journal of Hydrology. 2019; 575:1099-110.

[17] Chen W, Xie X, Wang J, Pradhan B, Hong H, Bui DT, et al. A comparative study of logistic model tree, random forest, and classification and regression tree

models for spatial prediction of landslide susceptibility. Catena. 2017; 151:147-60.
[18] Izquierdo-Verdiguier E, Zurita-Milla R. An evaluation of Guided Regularized Random Forest for classification and regression tasks in remote sensing. International Journal of Applied Earth Observation and Geoinformation. 2020; 88:102051.
[19] Hengl T, Nussbaum M, Wright MN, Heuvelink GB, Gräler B. Random forest as a generic framework for predictive modeling of spatial and spatio-temporal variables. PeerJ. 2018;6:e5518.
[20] Najafi-Vosough R, Faradmal J, Hosseini SK, Moghimbeigi A, Mahjub H. Predicting hospital readmission in heart failure patients in Iran: a comparison of various machine learning methods. Healthc Inform Res. 2021; 27(4):307-14.
[21] Breiman L. Random forests. Machine learning. 2001; 45(1):5-32.
[22] Hosseini-Esfahani F, Alafchi B, Cheraghi Z, Doosti-Irani A, Mirmiran P, Khalili D, et al. Using machine learning techniques to predict factors contributing to the incidence of metabolic syndrome in tehran: Cohort study. JMIR Public Health Surveill. 2021; 7(9):e27304.
[23] Pourhomayoun M, Shakibi M. Predicting mortality risk in patients with COVID-19 using machine learning to help medical decision-making. Smart Health. 2021; 20:100178.
[24] Cheng F-Y, Joshi H, Tandon P, Freeman R, Reich DL, Mazumdar M, et al. Using machine learning to predict ICU transfer in hospitalized COVID-19 patients. J Clin Med. 2020; 9(6):1668.
[25] Tezza F, Lorenzoni G, Azzolina D, Barbar S, Leone LAC, Gregori D. Predicting in-Hospital Mortality of Patients with COVID-19 Using Machine Learning Techniques. J Pers Med. 2021;11(5):343.
[26] Wollenstein-Betech S, Cassandras CG, Paschalidis IC. Personalized predictive models for symptomatic COVID-19 patients using basic preconditions: hospitalizations, mortality, and the need for an ICU or ventilator. Int J Med Inform. 2020; 142:104258.
[27] Das AK, Mishra S, Gopalan SS. Predicting CoVID-19 community mortality risk using machine learning and development of an online prognostic tool. PeerJ. 2020; 8:e10083.
[28] Wang J, Yu H, Hua Q, Jing S, Liu Z, Peng X, et al. A descriptive study of random forest algorithm for predicting COVID-19 patients outcome. PeerJ. 2020; 8:e9945.
[29] Kafieh R, Arian R, Saeedizadeh N, Amini Z, Serej ND, Minaee S, et al. Covid-19 in iran: Forecasting pandemic using deep learning. Computational and mathematical methods in medicine. 2021;2021.
[30] Mahdavi M, Choubdar H, Zabeh E, Rieder M, Safavi-Naeini S, Jobbagy Z, et al. A machine learning based exploration of COVID-19 mortality risk. Plos one. 2021;16(7):e0252384.
[31] Orooji A, Kazemi-Arpanahi H, Kaffashian M, Kalvandi G, Shanbehzadeh M. Comparing of Machine Learning Algorithms for Predicting ICU admission in COVID-19 hospitalized patients. Health Education and Health Promotion. 2021;9(3):229-36.
[32] Team RC. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria. URLhttps://www.R-project.org/. 2020.
[33] Zhao Z, Chen A, Hou W, Graham JM, Li H, Richman PS, et al. Prediction model and risk scores of ICU admission and mortality in COVID-19. PloS one. 2020;15(7):e0236618.
[34] Ji D, Zhang D, Xu J, Chen Z, Yang T, Zhao P, et al. Prediction for progression risk in patients with COVID-19 pneumonia: the CALL score. Clin Infect Dis. 2020; 71(6):1393-9.
[35] Lu J, Hu S, Fan R, Liu Z, Yin X, Wang Q, et al. ACP risk grade: a simple mortality index for patients with confirmed or suspected severe acute respiratory syndrome coronavirus 2 disease (COVID-19) during the early stage of outbreak in Wuhan, China. MedRxiv. 2020.
[36] Flythe JE, Assimon MM, Tugman MJ, Chang EH, Gupta S, Shah J, et al. Characteristics and outcomes of individuals with pre-existing kidney disease and COVID-19 admitted to intensive care units in the United States. Am J Kidney Dis. 2021; 77(2):190-203. e1.
[37] Buonamico E, Quaranta VN, Boniello E, Dimitri M, Di Lecce V, Labate L, et al. Risk factors for transfer from respiratory intermediate care unit to intensive care unit in COVID-19. Respir Investig. 2021; 59(5):602-7.
[38] Moulaei K, Shanbehzadeh M, Mohammadi-Taghiabad Z, Kazemi-Arpanahi H. Comparing machine learning algorithms for predicting COVID-19 mortality. BMC Med Inform Decis Mak. 2022; 22(1):1-12.
[39] Yadaw AS, Li Y-c, Bose S, Iyengar R, Bunyavanich S, Pandey G. Clinical features of COVID-19 mortality: development and validation of a clinical prediction model. Lancet Digit Health. 2020; 2(10):e516-e25.
[40] Gao Y, Cai G-Y, Fang W, Li H-Y, Wang S-Y, Chen L, et al. Machine learning based early warning system enables accurate mortality risk prediction for COVID-19. Nat Commun. 2020; 11(1):1-10.
[41] Shanbehzadeh M, Haghiri H, Afrash MR, Amraei M. Comparison of Machine Learning Tools for the Prediction of ICU Admission in COVID-19 Hospitalized Patients. Shiraz E Medical Journal. 2022.
[42] Lorenzen SS, Nielsen M, Jimenez-Solem E, Petersen TS, Perner A, Thorsen-Meyer H-C, et al. Using machine learning for predicting intensive care unit resource use during the COVID-19 pandemic in Denmark. Scientific reports. 2021; 11(1):1-10.