

Scientific Paper

Risk Factors Associated with In-Hospital Mortality in Iranian Patients with COVID-19: Application of Machine Learning

Sadjad SHAFIEKHANI^{1,ACDEF}, Sima RAFIEI^{2,ADEF,*}, Sina ABDOLLAHZADE^{3,4,AF}, Saber Souri^{5,BF}, Zeinab Moomeni^{5,BF}

¹Tehran University of Medical Science & Student's Scientific Research Center, Tehran University of Medical Science, Tehran, Iran

²Social Determinants of Health Research Center, Qazvin University of Medical Sciences, Qazvin, Iran

³Department of Neurosurgery, Royal Preston Hospital, Lancashire Teaching Hospitals NHS Trust, Preston, UK

⁴Department of Surgery, Qazvin University of Medical Sciences, Qazvin, Iran

⁵Student Research Center, School of Public Health, Qazvin University of Medical Sciences, Qazvin, Iran

*Corresponding author: Sima Rafiei; sima.rafie@gmail.com

(received 9 September 2021; revised 15 and 28 November 2021, and 4 January 2022; accepted 10 January 2022)

Abstract

Introduction: Predicting the mortality risk of COVID-19 patients based on patient's physiological conditions and demographic characteristics can help optimize resource consumption along with the provision of effective medical services for patients. In the current study, we aimed to develop several machine learning models to forecast the mortality risk in COVID-19 patients, evaluate their performance, and select the model with the highest predictive power.

Material and methods: We conducted a retrospective analysis of the records belonging to COVID-19 patients admitted to one of the main hospitals of Qazvin located in the northwest of Iran over 12 months period. We selected 29 variables for developing machine learning models incorporating demographic factors, physical symptoms, comorbidities, and laboratory test results. The outcome variable was mortality as a binary variable. Logistic regression analysis was conducted to identify risk factors of in-hospital death.

Results: In prediction of mortality, Ensemble demonstrated the maximum values of accuracy (0.8071, 95%CI: 0.7787, 0.8356), F1-score (0.8121 95%CI: 0.7900, 0.8341), and AUROC (0.8079, 95%CI: 0.7800, 0.8358). Including fourteen top-scored features identified by maximum relevance minimum redundancy algorithm into the subset of predictors of ensemble classifier such as BUN level, shortness of breath, seizure, disease history, fever, gender, body pain, WBC, diarrhea, sore throat, blood oxygen level, muscular pain, lack of taste and history of drug (medication) use are sufficient for this classifier to reach to its best predictive power for prediction of mortality risk of COVID-19 patients.

Conclusions: Study findings revealed that old age, lower oxygen saturation level, underlying medical conditions, shortness of breath, seizure, fever, sore throat, and body pain, besides serum BUN, WBC, and CRP levels, were significantly associated with increased mortality risk of COVID-19 patients. Machine learning algorithms can help healthcare systems by predicting and reduction of the mortality risk of COVID-19 patients.

Keywords: risk factors; machine learning; COVID-19; mortality.

Introduction

On 31 December, 2019, China reported several pneumonia cases of unknown etiology to the World Health Organization (WHO). Two months later, the WHO officially introduced the disease as an acute respiratory syndrome (SARS-CoV-2) caused by the COVID-19 virus. Since the new infectious disease appeared in Asia, the virus has spread across the globe, and later in March 2020, it turned into a global pandemic.¹

In a study conducted to develop a model of required and available beds for coronavirus patients, Zhuang et al. declared that the evolution of the outbreak could lead to the severe

shortage of hospital beds and medical resources in less than a month, putting an enormous strain on healthcare systems. They also confirmed that when hospitals were overwhelmed with COVID-19 patients, the mortality rate increased dramatically.² Several studies also addressed the issue and highlighted the scarcity of critical care equipment and hospital facilities in the face of the COVID-19 pandemic.^{3,4} Thus, it would be helpful to predict the prognosis of the disease precisely to lower the huge burden on healthcare systems and effectively allocate limited medical resources. To attain this goal, prediction models combine several factors to estimate the risk of dying or having

poor treatment outcomes in infected patients.⁵ Accordingly, Deschepper et al. affirmed that early prediction of mortality in hospitalized patients not only facilitates healthcare resources to be allocated more efficiently, it also enables healthcare providers to take immediate and decisive action to save the greatest number of lives.⁶ Bhattacharya also declared that identifying patients at greatest risk of mortality due to COVID-19 could lead to critical decisions such as tracking critical hospital resources leading to fair allocation of intensive care unit resources.⁷ Literature has recognized some of the demographic and clinical risk factors for severe COVID-19 progression and mortality.⁸ Older male adults with comorbidities, including hypertension, chronic obstructive pulmonary disease, cardiovascular disease, and obesity, were reported to be vulnerable to more severe morbidity and death after being infected with the COVID-19 virus.^{9,10} Most of the identified risk factors in recent research were related to the general population, not those admitted to hospitals with the disease. As several hematological and biochemical factors might also be associated with the adverse prognosis of the disease, there is a need to conduct similar studies among hospitalized patients.^{8,11}

Recently, researchers have used several machine learning models to accurately predict the severity of COVID-19 disease and its influencing factors. Most of them could estimate the probability of severe morbidity or mortality in the patients providing useful information for clinicians to manage infected individuals better.¹² Predicting the mortality risks based on patients' physiological conditions, symptoms, and demographic characteristics can help healthcare facilities prioritize those who are in urgent need to be hospitalized and receive critical care to reduce delays in the provision of health services to them.¹²

However, most of the models have used one kind of machine learning algorithm, and several studies have restricted the assessment of their models' performance to the area under the curve (AUC) analysis.¹³ In the current study, we aimed to develop a model consisting of a set of algorithms for processing data to explore new attributes, deal with missing data, remove redundant data factors, and choose the most useful ones. Following the data processing stage, we used machine learning algorithms to build a predictive model to predict whether a COVID-19 patient will die or survive. Furthermore, we aimed to develop several machine learning models, evaluate their performance, and select the model with the highest performance.

Materials and methods

We conducted a retrospective analysis of the records belonging to patients admitted to Bu-Ali-Sina hospital as one of the main hospitals of Qazvin located in the northwest of Iran over 12 months period (5 March 2020 to 30 March 2021). The entire 'Bu-Ali Sina' hospital is designated to provide healthcare services to COVID-19 patients.

Study population

Patients who were admitted and hospitalized with clinical symptoms of COVID-19 were included in our study. On the other hand, those who were admitted with non- COVID-19 symptoms or those with COVID-19 who were discharged from the emergency room without hospital admission were excluded. We divided the study population into two groups: the first group is related to patients who died (mean of age 64 years), containing 8.65% of individuals included in this study (34.5% female and 65.5% male), and the second group belongs to those who recovered (mean of age 55 years) with 61.34% of total population (43.7% female and 56.3% male).

Data collection

To identify the patients who met the inclusion criteria, two clinicians reviewed patient records of those who were hospitalized with documented positive COVID-19 test. After reviewing the literature and consulting with several physicians specialized in internal medicine, infectious diseases, and pulmonologist, we selected 29 variables for developing machine learning models. The variables incorporated demographic factors such as age and sex, hospitalization duration, blood oxygen level, physical symptoms such as fever, chill, sore throat, cough, panting breath, shortness breath, convulsion, weakness, stomachache, muscular pain, anorexia, diarrhea, body pain, headache, nausea, lack of smell and taste, history of drug (medication) use, history of the disease; laboratory test results including the level of WBC, HB, CRP, BUN, NA, and K. The outcome variable was mortality as a binary variable (yes if the patient died, and no if otherwise).

Data set preparation

We removed the patients with missing values or incomplete data features. Then, we divided COVID-19 patients into two groups. The imputed data were categorized into the train (90%) and test (10%) subsets and was used to train and test (validate) different classifiers, including Support Vector Machine (SVM), *K* nearest neighbor (KNN), Ensemble, Naïve Bayes and Decision tree.

Statistical analysis

We used t-test and chi-square analyses to examine which of the quantitative and categorical features were statistically different between the two groups. Also, logistic regression analysis was used to investigate the association of features with the probability of group 1 (die due to being infected by coronavirus) or group 2 (recover and survive). Statistical analyses were performed by MATLAB 2019b and IBM SPSS Statistics 19.

Machine learning models

We implemented different classifiers, including SVM, KNN, Naïve Bayes, Decision tree, and Ensemble models to predict the

mortality risk of COVID-19 in patients based on their demographic characteristics, symptoms, and experimental measurements. We trained classifiers by 10-fold random cross-validation in the training data set (90% of data set) and the trained models were validated by test data set (10% of data set). We performed hyperparameter optimization by Bayesian algorithm through training data set and optimized trained classifiers were tested by test data set. The values of model assessment metrics presented in the following sections are assessed in the test data set. Maximum Relevance Minimum Redundancy (MRMR) algorithm was also used to score features for classifying two groups of patients. MRMR algorithm selects a subset of features that have the minimum mutual information among themselves and have maximum mutual information by the target (output class). Different model assessment metrics, including accuracy, precision, recall, F1-score, specificity, and Area Under Receiver Operating Curve (AUROC), were computed in the different settings of features scored by the MRMR algorithm. These metrics (except AUROC) were calculated based on TP (True positive), TN (True negative), FP (False positive), and FN (False negative) values.

The ensemble classifier is constructed by the bagging (bootstrap aggregation) method. The weak learners of the ensemble model are tree classifiers that parallelly and independently evaluate the data samples, and final predictions of the ensemble model are determined by combining the predictions of weak learners. The MATLAB built-in function *fitensemble*, *fitsvm*, *fitcnb*, *fitctree* and *fitcknn* has been used to train ensemble, SVM, Naïve Bayes, decision tree, and KNN models, respectively. For more information about the hyperparameter of models, the file of the trained ensemble, SVM, tree, naïve Bayes, and KNN classifiers have been attached in an electronic supplementary file.

In the Naïve Bayes classification model, we used normal kernel functions for all of the features that their parameters have been optimized by the Bayesian optimization method.

In the decision tree classification model, Bayesian optimization was used to estimate the best values for the minimum value of leaf size, the maximum number of splits, and select the splitting criterion among Gini diversity index (GDI), deviance and twoing. The optimization revealed that the GDI and deviance were the best splitting criterion, and the optimum minimum number of parents and leaf nodes were 12 and 6, respectively.

In the KNN classification model, Bayesian optimization was used to estimate the number of neighbors, and select the best distance calculating method among 'city block', 'Chebyshev', 'euclidean', 'hamming', 'cosine', 'correlation', 'Jaccard', 'Mahalanobis', 'Seuclidean', and 'spearman', and also, select distance weight and singularity exponent.

Results

As provided in **Table 1**, the average age of COVID-19 patients in group 1 (patients who died) was 64.92 years (SD: 15.86), and in group 2 (recovered or survived patients) was 55.58 years (SD: 15.04), respectively (p-value < 0.05). Also, the values of average blood oxygen level, White Blood Cell (WBC), Blood Urea Nitrogen (BUN), and Sodium (NA) level in group 1 were 80.55 (SD: 14.83), 8.31 (SD: 4.20), 23.29 (SD: 12.30), 139.07 (SD: 8.95), respectively, while these values in group 2 were 87.82 (SD: 8.18), 6.72 (SD: 3.07), 18.11 (SD: 9.59), and 135.21 (SD: 23), respectively (p-value < 0.05).

The statistically significant differences in these features between the two groups were analyzed by t-test. As provided in **Table 2**, chi-square analysis revealed that there were statistically significant differences between the two groups in features including chills, sore throat, panting breath, shortness breath, weakness, muscular pain, anorexia, diarrhea, body pain, headache, lack of taste, and history of diseases (p-value < 0.05).

Table 1. Mean and standard deviations of features (with continuous values) of COVID-19 patients among two classes were compared, and their statistically significant differences were determined by t-test. The bolded italic lines represent a statistically significant (p-value < 0.05) difference.

Variables	Group 1 (dead)		Group 2 (recovered)		P-value
	Mean	SD	Mean	SD	
<i>Age</i>	64.92	15.86	55.58	15.04	2.47e-16
Hospitalization	7.15	4.71	7.25	4.60	0.7705
<i>Blood oxygen</i>	80.55	14.83	87.82	8.18	2.78e-18
<i>WBC</i>	8.31	4.20	6.72	3.07	9.38e-10
HB	13.35	2.11	13.35	2.27	0.9940
CRP	33.06	16.50	32.28	17.74	0.5341
<i>BUN</i>	23.29	12.30	18.11	9.59	4.91e-11
<i>NA</i>	139.07	8.95	135.21	23.00	0.0048
K	4.63	0.43	4.64	0.52	0.9601

Table 2. The results of Chi-square analysis to assess which features are significantly different between two groups of COVID-19 patients. The population and frequency rate of COVID-19 patients who have (and do not have) specific symptoms in each of the classes are provided.

Variables	Group 1 (dead)				Group 2 (recovered)				P-value
	Have symptom		Don't have symptom		Have symptom		Don't have symptom		
	Population	Freq. (%)	Population	Freq. (%)	Population	Freq. (%)	Population	Freq. (%)	
Fever	134	16.70	176	21.94	251	31.29	241	30.04	0.34
<i>Chills</i>	106	13.21	204	25.43	157	19.57	335	41.77	0.02
<i>Sore throat</i>	136	16.95	174	21.69	138	17.20	354	44.13	9.64e-06
Cough	237	29.55	73	9.10	349	43.51	143	17.83	0.11
<i>Panting breath</i>	251	31.29	59	7.35	360	44.88	132	16.45	0.03
<i>Shortness breath</i>	241	30.04	69	8.60	335	41.77	157	19.57	0.006
Convulsions	104	12.96	206	25.68	88	10.97	404	50.37	7.03
<i>Weakness</i>	161	20.07	149	18.57	207	25.81	285	35.53	0.03
Stomachache	47	5.86	263	32.79	97	12.09	395	49.25	0.84
<i>Muscular pain</i>	129	16.08	181	22.56	142	17.70	350	43.64	3.74e-09
<i>Anorexia</i>	110	13.71	200	24.93	124	15.46	368	45.88	9.72e-05
<i>Diarrhea</i>	45	5.61	265	33.04	113	14.08	379	47.25	0.01
<i>Body pain</i>	184	22.94	126	15.71	212	26.43	280	34.91	0.0006
<i>Headache</i>	116	14.46	194	24.18	169	21.07	323	40.27	0.001
Nausea	47	5.86	263	32.79	85	10.59	407	50.74	0.59
Lack of smell	78	9.725	232	28.92	125	15.58	367	45.76	0.66
<i>Lack of taste</i>	86	10.72	224	27.93	96	11.97	396	49.37	0.006
Drug (medication) use	149	18.57	161	20.07	152	18.95	340	42.39	0.42
<i>Disease history</i>	89	11.09	221	27.55	248	30.92	244	30.42	0.05

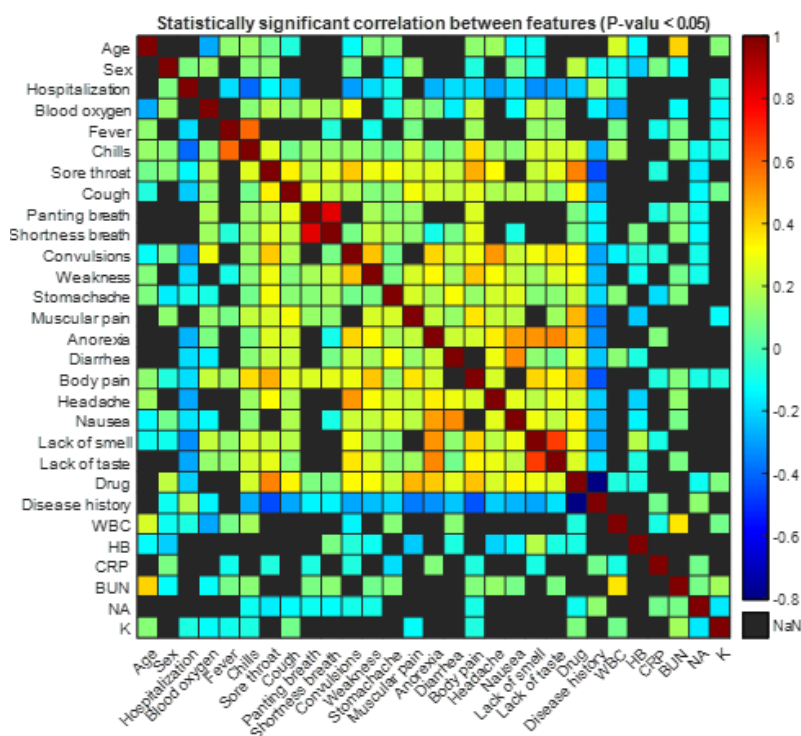


Figure 1. Statistically significant ($p\text{-value} < 0.05$) correlations between features. The black pixels represent there was no statistically significant correlation between features, and colored pixels show the strength of correlation in the range of $[-0.8, 1]$. Warm colors represent positive correlation, and cold colors describe the negative values of correlation between data set features.

The statistically significant correlations between data features are depicted in **Figure 1**. The black pixels ('NaN' or not a number) represent no significant correlation between features. We used Pearson and Spearman correlation for continuous (features presented in **Table 1**) and categorical (features mentioned in **Table 2**) features.

As shown in **Figure 1**, different clinical symptoms, including fever, chills, sore throat, cough, panting and shortness of breath, and so on., had a positive correlation with each other ($p\text{-value} < 0.05$). Also, the correlation between all features and outcome measures (mortality risk) is shown in **Table 3**.

Table 3. Statistically significant correlations (p-value < 0.05) between features and mortality risk.

Feature	Correlation value
Age	0.2839
Sex	-0.08838
Blood oxygen	-0.3012
Fever	0.07593
Sore throat	-0.1625
Panting breath	-0.08914
Shortness breath	-0.1045
Convulsion	-0.1787
Weakness	-0.09638
Muscular pain	-0.1313
Anorexia	-0.1101
Diarrhea	0.1035
Body pain	-0.1584
Lack of taste	-0.09568
History of drug use (medication)	-0.1727
Disease history	0.2061
WBC	0.2139
BUN	0.2294
NA	0.09939

As it is shown, the mortality risk has a positive correlation with age, fever, diarrhea, history of previous diseases, WBC, BUN, and NA (p-value < 0.05). The mortality risk inversely correlated with different features, including blood oxygen level, sore throat, panting and shortness of breath, etc., (p-value < 0.05).

$$\ln\left(\frac{P_{survive}}{P_{die}}\right) = -0.03X_{age} + 0.04X_{blood\ oxygen} - 0.63X_{stomachache} + 0.63X_{anorexia} + 0.59X_{body\ pain} - 1.90X_{lack\ of\ smell} + 1.35X_{lack\ of\ taste} - 0.43X_{disease\ history} - 0.06X_{WBC} - 0.01X_{CRP} - 0.02X_{BUN} + 0.37X_K \quad \text{Eq. 1}$$

Equation 1 was achieved by implementing the Logistic regression method to describe the effect of several features, including demographic, clinical symptoms, and experimental measurements on the response variable (mortality risk in COVID-19 patients). The coefficients of this equation describe statistically significant coefficients (p-value < 0.05).

The coefficients of the logistic regression model are provided in the heatmap of **Figure 2**. The first column of the figure includes the intercept term and relative risks estimated by the logistic regression model representing the ratio between the likelihood of surviving from COVID-19 disease versus the probability of dying. The second and third columns describe the lower and upper band of 95% confidence intervals of logistic regression coefficients. Furthermore, the fourth and fifth columns are depicting the ratio of probabilities $P_{survive}/P_{die}$ (odd ratio) and the corresponding p-values.

As shown in **Figure 2**, an increase in age by each one-unit increase in the age of COVID-19 patients increased the odds of death by 3% ($P_{die} = 1.0395 P_{survive}$). Our data analysis also revealed that a unit of increase in blood oxygen level increased the probability of surviving by 5% ($P_{survive} = 1.051 P_{die}$).

Furthermore, patients with a history of comorbidity were 1.5 times more likely to die ($P_{die} = 1.5399 P_{survive}$). Investigating the laboratory results revealed that a unit increase in white blood cell (WBC) count, Blood Urea Nitrogen (BUN), and C-reactive protein (CRP) reduced the likelihood of survival ($P_{survive} = 0.9355 P_{die}$), ($P_{survive} = 0.9706 P_{die}$), and ($P_{survive} = 0.9854 P_{die}$), respectively.

Figure 3 depicts the score of features for classifying two groups of COVID-19 patients (mortality versus recovering) determined by the MRMR algorithm. As shown in the figure, the six top-scored features were BUN, shortness of breath, seizure, fever, gender, and body pain.

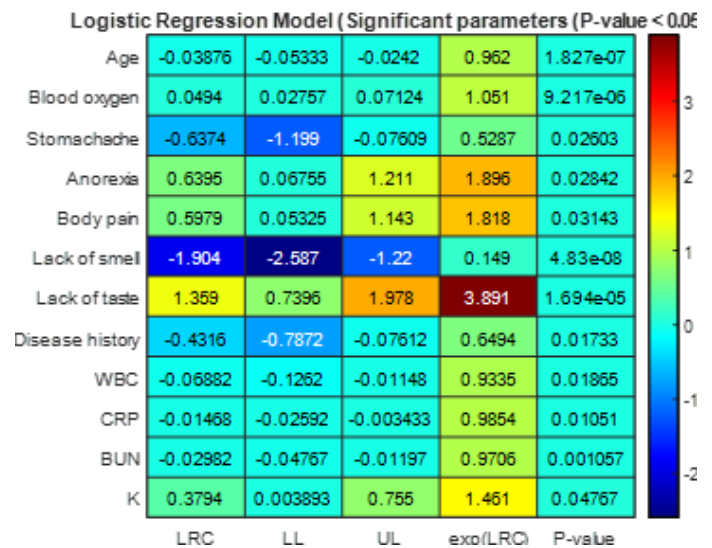


Figure 2. Logistic regression analysis of variables predicting the likelihood of surviving or dying due to being infected by a coronavirus. The values of logistic regression coefficients (LRC), lower level (LL), and upper levels (UL) of 95% confidence intervals of LRC values, e^{LRC} or odd ratio, and significant p-values (< 0.05) are shown in different columns of the figure.

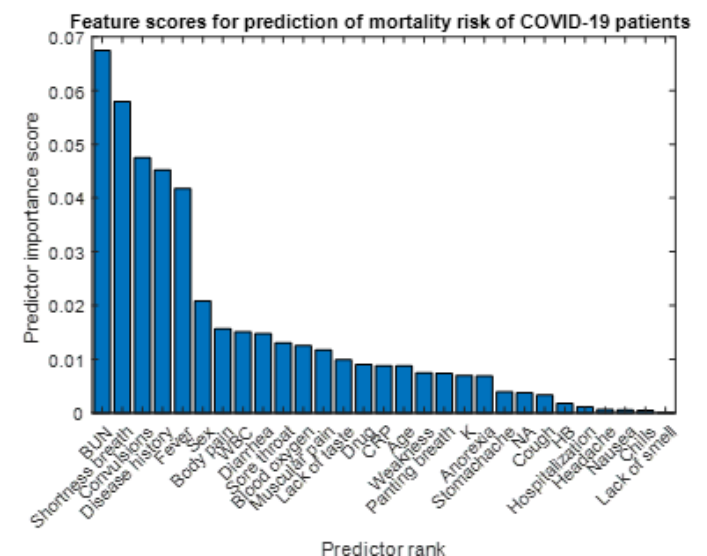


Figure 3. The MRMR algorithm identifies the predictors to distinguish two groups of COVID-19 patients.

In the next step, we assessed the predictive power of different classifiers, including SVM, KNN, Ensemble, DT, and NB to anticipate the mortality risk of COVID-19 patients. Accordingly, we tested the performance of these trained classification models by the validation data set. The values of 100 replicated average of model assessment metrics (along with the lower and upper level of 95% confidence intervals) in different subsets of top-scored features predicted by the MRMR algorithm are shown in **Figures 4-8**.

Actually, in each of the panels of **Figures 4-8**, the first red circle shows the computed classifier's assessment metric (the average among 100 replications) by including only the first top-scored feature (BUN) by MRMR into the feature space, and the second red circle represents the value of assessment metrics by including the first and second top scored features (BUN and shortness breath) into the feature space.

As shown in **Figure 4**, by including fourteen top-scored features identified by MRMR into the subset of predictors, the values of accuracy, precision, recall, F1-score, specificity, and AUROC of ensemble classifier led to its maximum values, and only this subset of features is sufficient for the ensemble to show its best predictive power. Adding more features to the feature space (after the fifteen top-scored features) led to a stable state and cannot increase the model assessment metrics.

Figure 5 depicts that the decision tree classifier had close efficiency with the ensemble model, except that the tree model needed fewer features (ten top-scored features identified by MRMR) to reach its best predictive power.

As shown in **Figures 6** (and **7**), the KNN (and SVM) models need 21 (and 20) top-scored features by MRMR to predict the mortality risk of COVID-19 patients by their maximum values of accuracy, F1-score, and AUROC. As shown in **Figure 8**, the maximum values of accuracy, F1-score, and AUROC of the Naïve Bayes model are obtained by considering the seventeen top-scored features by MRMR into the feature space.

As shown in **Figures 4-8**, all assessment metrics of classifiers did not show a stable pattern in response to an increase in the number of features added to their predictor subsets. Actually, in some cases, adding more features to the subset of predictors led to poor distinguishing power and reduction in model assessment metrics.

Comparing different classifiers revealed that ensemble and decision tree models had similar behavior and predictive power in the presence of different subsets of features scored by MRMR algorithms. Furthermore, NB and SVM had a similar trend in response to changing the subsets of predictors. The differences between these classifiers originate from different methods used to distinguish between two groups of patients.

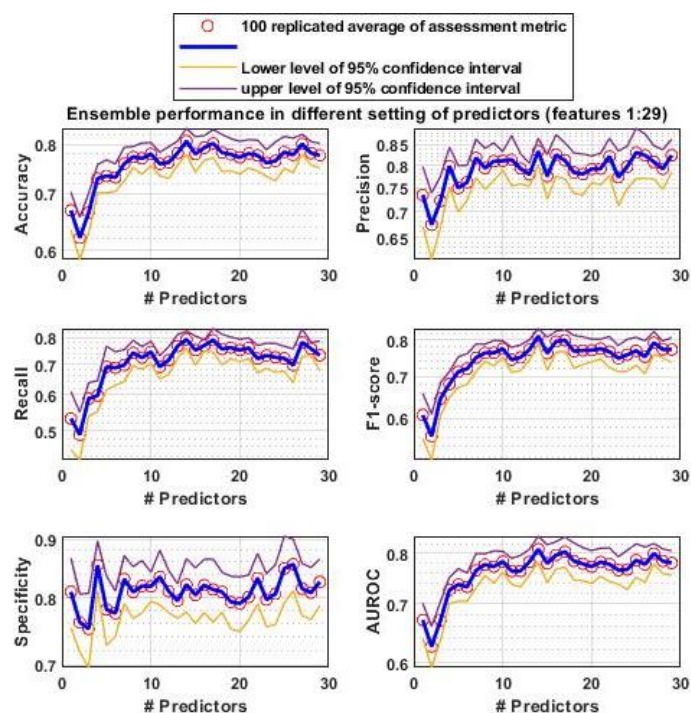


Figure 4. Ensemble classifier performance in the different settings of predictors scored by the MRMR algorithm (features 1:29). Different panels show the trend of accuracy (top-left panel), precision (top-right panel), recall (middle-left panel), F1-score (middle-right panel), specificity (down-left panel), and AUROC (down-right panel) in different settings of top scored features.

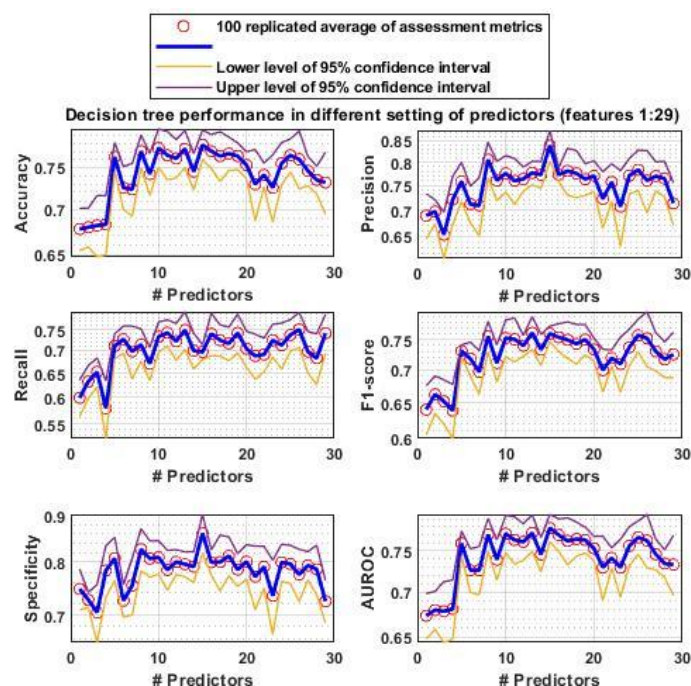


Figure 5. Decision tree classifier performance in different settings of predictors scored by MRMR algorithm (features 1:29)

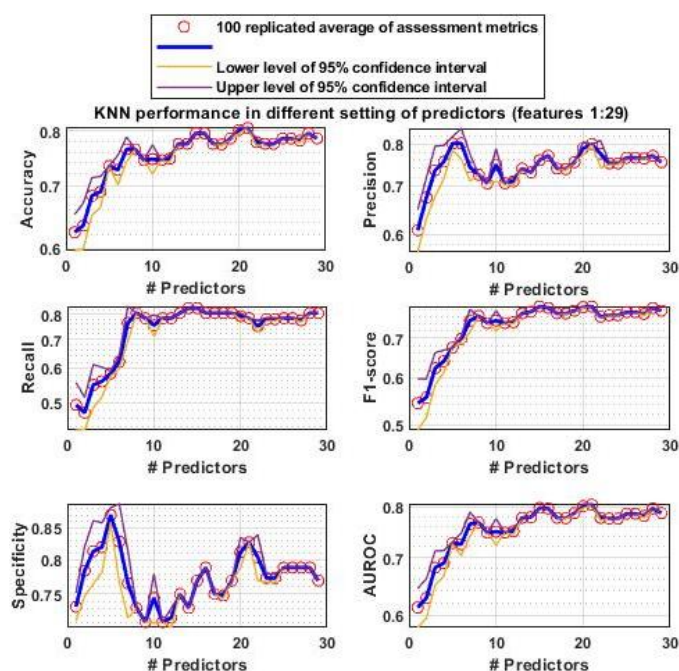


Figure 6. KNN classifier performance in different settings of predictors scored by MRMR algorithm (features 1:29)

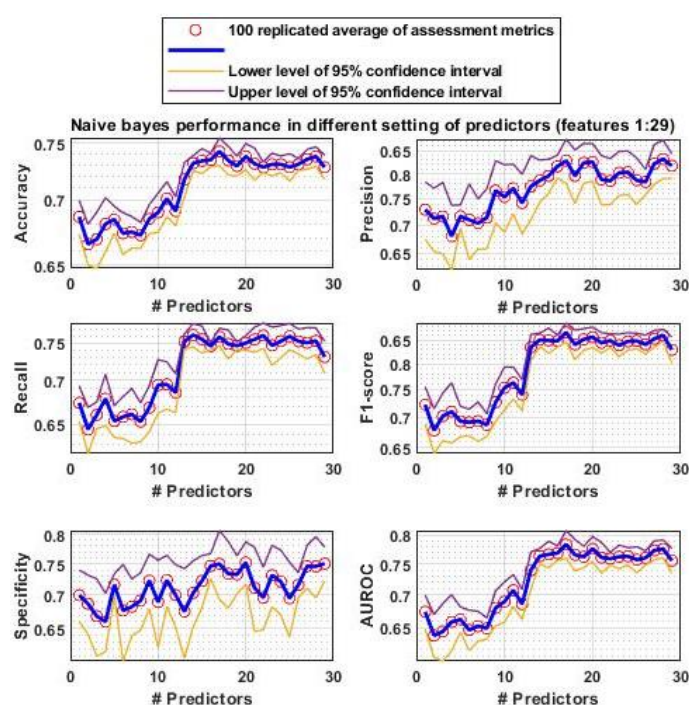


Figure 8. Naïve Bayes classifier performance in the different settings of predictors scored by MRMR algorithm (features 1:29)

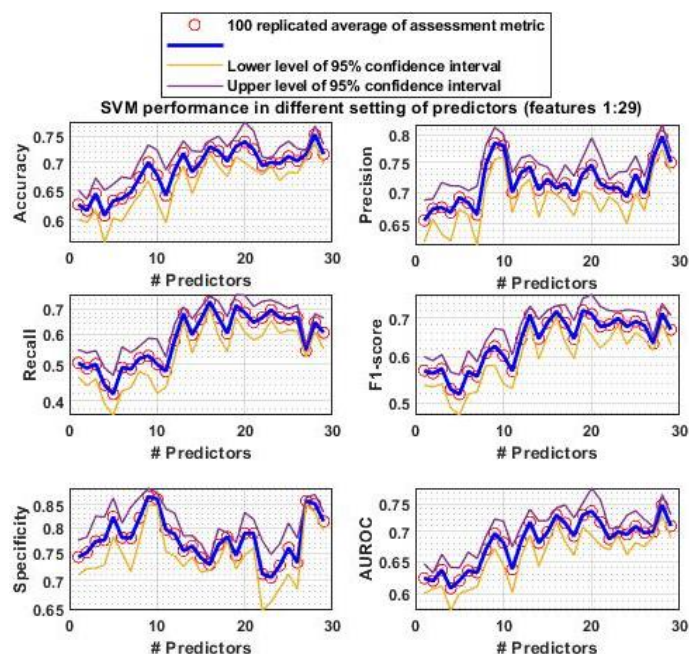


Figure 7. SVM classifier performance in different settings of predictors scored by MRMR algorithm (features 1:29)

Our data in **Table 4** and **Figures 4-8** revealed that ensemble model has the maximum values of accuracy (0.8071, 95%CI: 0.7787, 0.8356), F1-score (0.8121 95%CI: 0.7900, 0.8341), and AUROC (0.8079, 95%CI: 0.7800, 0.8358). Therefore, the ensemble had the best predictive power for mortality risk assessment among all classifiers implemented in this study.

Discussion

Our study findings revealed that machine learning models could be used to predict the mortality risk among COVID-19 patients using their socio-demographic characteristics, clinical symptoms, and medical history. We presented the results of the logistic regression model and showed the input variables which were significantly associated with mortality risk in COVID-19 patients. Similar to previous literature, it was found that old age, lower level of saturation, underlying medical conditions, symptoms including shortness of breath, seizure, fever, sore throat, and body pain, besides serum BUN, WBC, and CRP levels were significantly associated with mortality risk.^{8,12,13}

In line with recent research, our results depicted that patients older than 70 years were at higher risk of fatality due to COVID-19 disease. The prevalence of underlying medical conditions such as hypertension, diabetes mellitus, respiratory, cardiovascular, and chronic renal diseases in elderly people might be an important reason for their increased risk of mortality.^{10,16}

World Health Organization declared that the case-fatality rate (CFR) was higher in patients with diabetes, hypertension, cancer, cardiovascular and pulmonary disease.¹⁷ Literature also affirmed that the risk of severe COVID-19 leading to in-hospital mortality increases when a person has several underlying medical conditions. The most common comorbidities were reported to be hypertension, abnormally high levels of lipids in the blood, diabetes, ischemic heart disease, and chronic pulmonary diseases.^{18,19}

Table 4. Comparison assessment metrics of different classifiers implemented for prediction of mortality risk of COVID-19 patients. The values of 10 replicated average of classifiers assessment metrics along with their 95% confidence intervals.

Classifier	Accuracy	Precision	Recall	F1-score	Specificity	AUROC
Ensemble	0.8071 (0.7787, 0.8356)	0.8341 (0.8021, 0.8661)	0.7937 (0.7598, 0.8276)	0.8121 (0.7900, 0.8341)	0.8561 (0.8119, 0.9004)	0.8079 (0.7800, 0.8358)
SVM	0.7541 (0.7364, 0.7718)	0.7969 (0.7744, 0.8194)	0.7299 (0.7056, 0.7542)	0.7204 (0.6862, 0.7546)	0.8673 (0.8484, 0.8862)	0.7477 (0.7283, 0.7671)
Decision tree	0.7786 (0.7597, 0.7994)	0.8381 (0.8019, 0.8743)	0.7501 (0.7069, 0.7932)	0.7609 (0.7328, 0.7890)	0.8605 (0.8183, 0.9027)	0.7786 (0.7611, 0.7961)
Naïve bayes	0.7847 (0.7591, 0.8103)	0.8316 (0.7902, 0.8731)	0.7234 (0.6886, 0.7581)	0.7666 (0.7454, 0.7878)	0.8523 (0.8173, 0.8873)	0.7839 (0.7591, 0.8086)
KNN	0.8061 (0.8052, 0.8071)	0.8020 (0.7628, 0.8412)	0.8261 (0.8253, 0.8272)	0.7917 (0.7912, 0.7921)	0.8707 (0.8587, 0.8827)	0.8048 (0.8024, 0.8053)

The relationship between kidney failure and fatal in-hospital outcomes among patients with Corona Virus has also been confirmed in previous studies.^{20,21}

A study conducted by Mazinani and Rude (2020) revealed that oxygen saturation and respiratory symptoms such as sore throat and shortness of breath were important markers for differentiating disease severity.²² Several pieces of research affirmed the existence of a strong relationship between hypoxia, hospitalization, and death in COVID-19 patients, introducing oxygen saturation below 90% as a strong risk factor for mortality outcomes.^{10,23} Xie et al added that as hypoxemia was proven to be related to an increased likelihood of inflammation, infectious markers including WBC counts, D-dimer, and CRP levels also showed an increasing level.²⁴ Symptoms including shortness of breath, body pain, and fatigue were also markers of disease exacerbation and mortality.²⁵ Accordingly, Chen et al. found that dyspnea was associated with a four-fold increase in the mortality risk highlighting the need for early recognition of patients with poor outcomes.⁹ Liang et al. also revealed that decreased consciousness and difficulty in breathing were among key predictive factors of the disease severity and its related mortality.²⁶

Cough as a symptom of infection in the respiratory tract was found to be associated with severe COVID-19. In similar studies, exhaustion, spitting out sputum, and stuffy nose were mentioned as common predictive symptoms of disease exacerbation and poor outcome.²⁷ Studies also mentioned sustained fever as a risk factor for poor outcomes. For example, in research conducted by Guan et al. fever was significantly associated with the severity of COVID-19 disease. They added that patients who experienced worsening disease conditions during hospitalization were among those with persistent fever.²⁰

In our study, laboratory analysis revealed that elevated levels of BUN, CRP, and WBC were potential risk factors for worse outcomes in COVID-19 patients. This finding is consistent with the results of Lin et al. study which reported higher values of WBC, BUN, and CRP in severe COVID-19 patients.²⁸ Medical evidence also mentioned BUN as an end product of protein metabolism which was associated with fatality in different illnesses.^{29,30} A study conducted by Wernly et al. highlighted the

predictability role of BUN in the mortality of patients admitted to ICU.³¹ Similarly, Aronson et al. found that elevated BUN could act as a risk factor for fatality in patients admitted for heart failure.³² Literature also mentioned the ratio of BUN to serum albumin as a key predictive factor of disease severity and mortality in patients with pneumonia.^{33,34} CRP was another risk factor connected to COVID-19 poor outcomes and in-hospital mortality.³⁵ In a study by Sharifpour et al., the efficacy of CRP values in the prediction of COVID-19 severity and fatality was agreed. Their findings revealed that the level of CRP was significantly higher in patients who died, compared to those who recovered.³⁶ Similarly, research conducted among COVID-19 patients found that CRP level in patients who died as a result of COVID-19 during hospitalization was 10-fold higher than survivors.²⁶ Furthermore, we found that the number of white blood cells was significantly higher in the group of patients with adverse outcomes. In line with our finding, Cekerevac et al. revealed a significant relationship between WBC count and the incidence of inflammation and thrombotic response.³⁷

In a study, the authors presented machine learning classification models including SVM, logistic regression, neural network, and gradient boosted decision tree to predict mortality risk of COVID-19 patients up to 20 days in advance. They used the lasso feature selection method, and their best value of AUROC in the validation group was 0.9760 (0.9613–0.9906).³⁸ Also, in another study with the same aim, the authors trained machine learning classification models to predict the mortality risk of COVID-19 patients on average 10 days ahead with 90% accuracy. Their models were trained by data of 375 COVID-19 patients gathered in China. They recorded three biomarkers, including lymphocyte, lactic dehydrogenase (LDH), and high-sensitivity C-reactive protein (hs-CRP).³⁹ In a study, the authors used clinical and epidemiological data of 183 severe COVID-19 patients (115 survivors and 68 nonsurvivors) to construct the predictive models. They trained different classification models, including elastic net, logistic regression, random forest, and partial least squares regression model, and their best value for AUROC was 88%.⁴⁰ In another study, the authors included 398 COVID-19 patients to develop machine learning models for the prediction of mortality risk of COVID-19 patients up to 2 days

before expiration. Their predictive features were blood urea nitrogen, c-reactive protein, serum albumin, serum calcium, and lactic. Their data analysis revealed that the SVM model achieved 93% AUROC for predicting patient mortality risk.⁴¹ In this study, we gathered demographic, clinical, and laboratory data of 802 COVID-19 patients to develop machine learning predictive models. We trained SVM, Tree, Ensemble, Naïve Bayes, and KNN classification models to predict whether a COVID-19 patient will die or survive. Our data analysis revealed that the Ensemble model had the best predictive power and its AUROC in the validation group was 80%. Literature has affirmed that the Ensemble method can improve the accuracy of base models such as artificial neural networks, decision trees, and Naïve Bayes. Furthermore, one of the advantages of the Ensemble method is in handling a large volume of data. Thus in this study, as we had a large data set, this ability could play a beneficial role in conducting the simulations and gave better accuracy than other methods for real-world problems.^{42,43}

This study has several limitations. First, the data used in this study did not encompass diagnostic information regarding radiologic results, which might act as significant predictive factors. Second, the limitations of retrospective observational studies conducted in a single center might restrict the generalization of findings and give a narrow-minded view of the general circumstances. Third, the effect of different treatment types was not taken into consideration as one of the confounding factors; however, as all COVID-19 patients are referred to

designated hospitals in Qazvin with standard therapy, it is assumed that the study population in our study was treated with the same type of treatment.

Conclusion

Due to an increasing number of patients and limited hospital resources, early diagnosis of patients with a critical medical condition is essential to ensure the provision of proper medical care with efficient use of available resources. In conclusion, we developed machine learning models to predict the risk of mortality in hospitalized COVID-19 patients. The ensemble classifier had the best predictive power for mortality risk assessment among all other classifiers implemented in the study. This model had the maximum values of accuracy, F1-score, and AUROC. Age, level of oxygen saturation, underlying medical conditions, shortness of breath, seizure, fever, sore throat, and body pain, besides serum BUN, WBC, and CRP levels were identified as predictors of COVID-19 mortality risk.

Acknowledgment

We would like to appreciate the efforts of doctors, nurses, and all healthcare workers trying to keep us safe in the COVID-19 crisis.

References

1. World Health Organization. Pneumonia of unknown cause 2020, 5 January [Available from: <https://www.who.int/csr/don/05-january-2020-pneumonia-of-unknown-cause-china/en/>. Accessed 8 June 2020.
2. Zhuang Z, Cao P, Zhao S, Han L, He D, Yang L. The shortage of hospital beds for COVID-19 and nonCOVID-19 patients during the lockdown of Wuhan, China. *Ann Transl Med* 2021;9(3):200. <https://doi.org/10.21037/atm-20-5248>
3. Li J, Yuan P, Heffernan J, et al. Observation wards and control of the transmission of COVID-19 in Wuhan. *Bull World Health Organ* 2020.
4. Sen-Crowe B, Sutherland M, McKenney M, Elkbuli A. A Closer Look in to Global Hospital Beds Capacity and Resource Shortages During the COVID-19 Pandemic. *Journal of Surgical Research* 2021;260:P53-63. <https://doi.org/10.1016/j.jss.2020.11.062>
5. Gerayelia FV, Milne S, Cheunga Ch, Lia X, Tony Yanga Ch. W, Tama A, Choia L.H, Baea A, Sin D.D. COPD and the risk of poor outcomes in COVID-19: A systematic review and meta-analysis. *E Clinical Medicine* 2021;33:100789. <https://doi.org/10.1016/j.eclinm.2021.100789>
6. Deschepper M, Waegeman W, Vogelaers D, Eeckloo K. Using structured pathology data to predict hospital-wide mortality at admission. *PLoS One*. 2020;15(6):e0235117. <https://doi.org/10.1371/journal.pone.0235117>
7. Bhattacharya S, Rajan V, Shrivastava H. ICU mortality prediction: a classification algorithm for imbalanced datasets. *Proceedings of the Thirty-First AAAI Conference on Artificial Intelligence*. 2017;31(1):1288-94.
8. Tian, W. et al. Predictors of mortality in hospitalized COVID-19 patients: a systematic review and meta-analysis. *J. Med. Virol.* 2020;92(10):1875-1882. <https://doi.org/10.1002/jmv.26050>
9. Chen T, Wu D, Chen HL, Yan WM, Yang DL, Chen G, et al. Clinical characteristics of 113 deceased patients with coronavirus disease 2019: retrospective study. *BMJ*. 2020; 368:m1091. <https://doi.org/10.1136/bmj.m1091>
10. Zhou F, Yu t, Du R, et al. Clinical course and risk factors for mortality of adult in patients with COVID-19 in Wuhan, China: a retrospective cohort study. *Lancet*. 2020;395(10229):1054-1062. [https://doi.org/10.1016/S0140-6736\(20\)30566-3](https://doi.org/10.1016/S0140-6736(20)30566-3)

11. Bikdeli B. et al. COVID-19 and thrombotic or thromboembolic disease: implications for prevention, antithrombotic therapy, and follow-up. *J. Am. Coll. Cardiol.* 2020;75(23):2950-2973. <https://doi.org/10.1016/j.jacc.2020.04.031>
12. 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. <https://doi.org/10.1016/j.smhl.2020.100178>
13. Alballa N, Al-Turaiki I. Machine learning approaches in COVID-19 diagnosis, mortality, and severity risk prediction: A review. *Informatics in Medicine Unlocked.* 2021;24(100564). <https://doi.org/10.1016/j.imu.2021.100564>
14. Gong J. et al. A tool to early predict severe corona virus disease 2019 (COVID-19): A multicenter study using the risk nomogram in Wuhan and Guangdong, China. *Clin. Infect. Dis.* 2020;71(15):833-840. <https://doi.org/10.1093/cid/ciaa443>
15. Yuan, M., Yin, W., Tao, Z., Tan, W. & Hu, Y. Association of radiologic findings with mortality of patients infected with 2019 novel coronavirus in Wuhan, China. *PLoS ONE.* 2020;15(3):e0230548. <https://doi.org/10.1371/journal.pone.0230548>
16. Wang L, He W, Yu X, et al. Coronavirus Disease 2019 in elderly patients: Characteristics and prognostic factors based on 4-week follow-up. *J. Infect.* 2020;80:639-645. <https://doi.org/10.1016/j.jinf.2020.03.019>
17. WHO Coronavirus disease (COVID-2019) situation reports (2020).
18. De Giorgi A, F. Fabbian, S. Greco, et al. Prediction of in-hospital mortality of patients with SARS-CoV-2 infection by comorbidity indexes: an Italian internal medicine single center study. *Eur Rev Med Pharmacol Sci.* 2020;24(19):10258-10266. https://doi.org/10.26355/eurrev_202010_23250
19. Dominguez-Ramirez L, Rodriguez-Perez F, Sosa-Jurado F, et al. The role of metabolic comorbidity in COVID-19 mortality of middle-aged adults. The case of Mexico. *medRxiv* 2020.12.15.20244160. <https://doi.org/10.1101/2020.12.15.20244160>
20. Guan WJ, Liang WH, Zhao Y, Liang HR, Chen ZS, Li YM, et al. Comorbidity and its impact on 1590 patients with COVID-19 in China: a nationwide analysis. *Eur Respir J.* 2020;55:2000547. <https://doi.org/10.1183/13993003.01227-2020>
21. Cheng Y, Luo R, Wang K, Zhang M, Wang Z, Dong L, et al. Kidney disease is associated with in-hospital death of patients with COVID-19. *Kidney Int.* 2020;97:829-838, <https://doi.org/10.1016/j.kint.2020.03.005>
22. Mazinani M., Rude B.J. The novel zoonotic coronavirus disease 2019 (COVID-19) pandemic: Health perspective on the outbreak. *J Healthc Qual Res.* 2020;36(1):47-51. <https://doi.org/10.1016/j.jhqr.2020.09.004>
23. Wu C, Chen X, Cai Y, et al. Risk factors associated with acute respiratory distress syndrome and death in patients with coronavirus disease 2019 pneumonia in Wuhan, China. *JAMA Intern Med.* 2020;180(7):934-943. <https://doi.org/10.1001/jamainternmed.2020.0994>
24. Xie J, Covassin N, Fan Zh, Singh P, Gao W, Li G, et al. Association between Hypoxemia and Mortality in Patients With COVID-19. *Mayo Clin Proc* 2020;95(6):1138-1147. <https://doi.org/10.1016/j.mayocp.2020.04.006>
25. Xiang G, Xie L, Chen Zh, Hao Sh, Fu C, Wu Q, Liu X, Li Sh. Clinical risk factors for mortality of hospitalized patients with COVID-19: systematic review and meta-analysis. *Annals of Palliative Medicine.* 2021;10(3). <https://doi.org/10.21037/apm-20-1278>
26. Liang W, Liang H, Ou L, Chen B, Chen A, Li C, et al. China Medical Treatment Expert Group for COVID-19. Development and validation of a clinical risk score to predict the occurrence of critical illness in hospitalized patients with COVID-19. *JAMA Intern Med.* 2020;180:1081-1089. <https://doi.org/10.1001/jamainternmed.2020.2033>
27. Wu JT, Leung K, Bushman M, Kishore N, Niehus R, de Salazar PM, et al. Estimating clinical severity of COVID-19 from the transmission dynamics in Wuhan, China. *Nat Med.* 2020;26:506-10. <https://doi.org/10.1038/s41591-020-0822-7>
28. Lin L, Lu L, Cao W, et al. Hypothesis for potential pathogenesis of SARS-CoV-2 infection-a review of immune changes in patients with viral pneumonia. *Emerg Microbes Infect.* 2020;9:1-14. <https://doi.org/10.1080/22221751.2020.1746199>
29. Cheng A, Hu L, Wang Y, Huang L, Zhao L, Zhang C et al. Diagnostic performance of initial blood urea nitrogen combined with D-dimer levels for predicting in-hospital mortality in COVID-19 patients. *Int J Antimicrob Agents.* 2020;56(3):106110. <https://doi.org/10.1016/j.ijantimicag.2020.106110>
30. Yang CJ, Chen J, Phillips AR, Windsor JA, Petrov MS. Predictors of severe and critical acute pancreatitis: a systematic review. *Dig Liver Dis.* 2014;46:446-451. <https://doi.org/10.1016/j.dld.2014.01.158>
31. Wernly B, Lichtenauer M, Vellinga NAR, Boerma EC, Ince C, Kelm M. Blood urea nitrogen (BUN) independently predicts mortality in critically ill patients admitted to ICU: a multicenter study. *Clin Hemorheol Microcirc.* 2018;69:123-131. <https://doi.org/10.3233/CH-189111>
32. Aronson D, Mittleman MA, Burger AJ. Elevated blood urea nitrogen level as a predictor of mortality in patients admitted for decompensated heart failure. *Am J Med.* 2004;116:466-473. <https://doi.org/10.1016/j.amjmed.2003.11.014>
33. Tokgoz Akyil F, Yalcinsoy M, Hazar A, Cilli A, Celenk B, Kilic O. Prognosis of hospitalized patients with community-acquired pneumonia. *Pulmonology.* 2018;24(3):164-169. <https://doi.org/10.1016/j.rppnen.2017.07.010>
34. Ryu S, Oh SK, Cho SU, You Y, Park JS, Min JH. Utility of the blood urea nitrogen to serum albumin ratio as a prognostic factor of mortality in aspiration pneumonia patients. *Am J Emerg Med.* 2021;43:175-179. <https://doi.org/10.1016/j.ajem.2020.02.045>

35. Chalmers JD, Singanayagam A, Hill AT. C-reactive protein is an independent predictor of severity in community-acquired pneumonia. *Am J Med.* 2008;121:219-225. <https://doi.org/10.1016/j.amjmed.2007.10.033>
36. Sharifpour M, Rangaraju S, Liu M, Alabyad D, Nahab FB, Creel-Bulos CM, et al. C-Reactive protein as a prognostic indicator in hospitalized patients with COVID-19. *PLoS ONE.* 2020;15(11):e0242400. <https://doi.org/10.1371/journal.pone.0242400>
37. Cekerevac I, Turnic TN, Dragicin N, Andjic M, Zivkovic V, Simovic S, et al. Predicting Severity and Intrahospital Mortality in COVID-19: The Place and Role of Oxidative Stress. *Oxidative Medicine and Cellular Longevity.* 2021:6615787. <https://doi.org/10.1155/2021/6615787>
38. Gao Y, Cay GY, Fang W, et al. Machine learning based early warning system enables accurate mortality risk prediction for COVID-19. *Nature communications.* 2020;11:5033. <https://doi.org/10.1038/s41467-020-18684-2>
39. Yan L, Zhang HT, Goncalves J, et al. n interpretable mortality prediction model for COVID-19 patients. *Nature Machine Intelligence.* 2020;5(2):283-288. <https://doi.org/10.1038/s42256-020-0180-7>
40. Hu C, Liu Z, Jiang Y, et al. Early prediction of mortality risk among severe COVID-19 patients using machine learning. *International Journal of Epidemiology.* 2020;49(6):1918–1929. <https://doi.org/10.1093/ije/dyaa171>
41. Booth AL, Abels E, McCaffrey P. Development of a prognostic model for mortality in COVID-19 infection using machine learning. *Modern Pathology.* 2021;34:522-531. <https://doi.org/10.1038/s41379-020-00700-x>
42. Kim Y. Boosting and measuring the performance of ensembles for a successful database marketing. *Expert Systems with Applications.* 2009;36:2161-76. <https://doi.org/10.1016/j.eswa.2007.12.036>
43. Piao Y, Park HW, Jin CH, Ryu KH. Ensemble method for classification of high-dimensional data. *International Conference on Big Data and Smart Computing (BIGCOMP).* 2014:245-249. <https://doi.org/10.1109/BIGCOMP.2014.6741445>