Received: 8 December 2022 | Revised: 20 February 2023 | Accepted: 20 March 2023

DOI: 10.1002/hsr2.1162

ORIGINAL RESEARCH

Health Science Reports Open Access WILEY

Predicting the mortality of patients with Covid-19: A machine learning approach

Hassan Emami[1] | Reza Rabiei[1] | Solmaz Sohrabei[1] | Alireza Atashi[2]

[1]Department of Health Information Technology and Management, School of Allied Medical Sciences, Shahid Beheshti University of Medical Sciences, Tehran, Iran

[2]Virtual School, Tehran University of Medical Sciences, Tehran, Iran

Correspondence
Reza Rabiei and Solmaz Sohrabei, Department of Health Information Technology and Management, School of Allied Medical Sciences, Shahid Beheshti University of Medical Sciences, Tehran, Iran.
Email: R.Rabiei@sbmu.ac.ir and solmazsohrabee1@gmail.com

Abstract

Background and Aims: Infection with Covid-19 disease can lead to mortality in a short time. Early prediction of the mortality during an epidemic disease can save patients' lives through taking timely and necessary care interventions. Therefore, predicting the mortality of patients with Covid-19 using machine learning techniques can be effective in reducing mortality rate in Covid-19. The aim of this study is to compare four machine-learning algorithm for predicting mortality in Covid-19 disease.

Methods: The data of this study were collected from hospitalized patients with COVID-19 in five hospitals settings in Tehran (Iran). Database contained 4120 records, about 25% of which belonged to patients who died due to Covid-19. Each record contained 38 variables. Four machine-learning techniques, including random forest (RF), regression logistic (RL), gradient boosting tree (GBT), and support vector machine (SVM) were used in modeling.

Results: GBT model presented higher performance compared to other models (accuracy 70%, sensitivity 77%, specificity 69%, and the ROC area under the curve 0.857). RF, RL, and SVM models with the ROC area under curve 0.836, 0.818, and 0.794 were in the second and third places.

Conclusion: Considering the combination of multiple influential factors affecting death Covid-19 can help in early prediction and providing a better care plan. In addition, using different modeling on data can be useful for physician in providing appropriate care.

KEYWORDS

Covid-19, gradient boosting tree, machine learning, random forest, support vector machine

1 | INTRODUCTION

The Covid-19 pandemic put pressure on the economic and health infrastructures all over the world. A greater focus on early clinical interventions could be beneficial in reducing mortality rate.[1] Patients with an acute condition need to be admitted to the intensive care unit (ICU) and use a ventilator. Studies have shown in that at the peak of this disease, about half of the patients with an acute condition in developed and developing countries were not able to receive special ICU equipment on time, and 30% of the deceased patients did not receive mechanical ventilation on time.[2] A sharp increase in the demand for medical resources, as well as a shortage of hospital beds and critical care equipment to treat patients in a timely manner can become a crisis; therefore, the rapid identification of patients at high risk of mortality can be a key factor in reducing mortality rate. The body's immune system is effective in improving the treatment process of viral diseases, but in this emerging disease, physicians are often unable to accurately predict the condition of patients with immune system defects affected by Covid-19 from the time of admission to the next stages of the disease, because this disease affects the immune system of each person differently. In addition, due to the nature of the Covid-19, a patient's stable condition can rapidly deteriorate[3] and challenge even the most skilled physicians. Data from epidemiological studies show that severe disease occurs in approximately 20% of patients and at older ages. In addition, comorbidities and cardiovascular diseases are associated with a worse prognosis.[3,4] Therefore, identifying key factors in patients with Covid-19 is critical in the prognosis of the disease.[5,6] Early prediction of mortality rate among patients with Covid-19 can lead to efficient allocation of resources and setting effective care plans. In recent years, there has been an increasing attention to the application of machine learning in predicting diseases through identifying complicated patterns in large data sets.[7,8] This study aimed to apply machine-learning approaches in predicting the mortality of patients with Covid-19. Studies have shown that elderly patients have a weaker immune response to Covid-19, therefore, the elderly have also been a group of the population that has been severely affected by the infection of Covid-19.[9,10] Also, it has been reported that mortality is higher in men than women, which could be due to reasons including higher prevalence of chronic diseases among men (cardiovascular disease, hypertension, and lung disease) as well as higher smoking compared to women. Other studies have explained the gender difference in Covid-19 mortality with higher ACE2 (angiotensin-converting enzyme 2) receptor levels observed in Asian men[11–13]

To identify influential risk factors for in-hospital death due to Covid-19, the OpenSAFELY platform investigated the 17.4 million of British patients electronic health records. Scientists have identified the number of risk factors affecting in mortality of Covid-19 such as old age, socioeconomic status, sex, race, and some clinical conditions (hypertension, diabetes, obesity, cancer, respiratory diseases, heart, kidney, liver, neurological, and autoimmune conditions).[14–16]

In 93 countries, to the correlation between the presence of the disease and the prevalence of death in patients with Covid-19, eight diseases (asthma, lung cancer, chronic obstructive pulmonary disease [COPD], Alzheimer's disease [AD], high blood pressure, ischemic heart disease, depression, and diabetes) were analyzed. Also, the analysis of six social and demographic factors [unemployment, age over 65, urbanization, population density, and social and demographic index] showed that the case mortality rates of Covid-19 in countries with a high prevalence of AD, lung cancer, asthma, and COPD risk factors were more, comorbidities such as AD and lung diseases that may be more influential than elderly alone[17,18] In the mentioned studies, the risk of mortality among Covid-19 patients with BMI > 30 were 118% higher compared to patients with normal weight. In addition, patients with a history of smoking were 81% more at risk of death compare to patients without a history of smoking. The prevalence of mortality in Covid-19 patients hospitalized in the ICU was very high. The mortality rate of patients hospitalized in the ICU was observed 272% higher than non-ICU patients. The capacity of ICU beds in developing countries are limited, this limitation may be the main reason for high level of mortality patients.[19,20] To identify the clinical, laboratory, and demographic factors affecting the Covid-19, after interviews with Infectious disease doctors, specialized papers published in this field have been reviewed and after collecting relevant data, those have been processed for modeling. The influencing risk factors in mortality Covid-19 are shown in Table 1.

2 | METHODS

2.1 | Samples definition

This study was conducted on the data of patients admitted to five hospitals affiliated to Tehran Medical Sciences (TUMS) between 2020 and 2021. The data set included 4140 records and 38 variables. In the medical centers after admission for each patient with Covid-19, a series of diagnostic tests, physical examinations such as height, weight, temperature, blood pressure, systolic/diastolic pressure, respiration, number, and heart rate were performed. Chest imaging and blood concentration tests were performed to control the progress of the disease and prevent stroke.

2.2 | Preprocessing

Outlier data records were removed from the data set and missing data were replaced by K-NN technique ($K = 30$ and 50). To balance the data, the SMOTE (Synthetic Minority Over-sampling Technique) technique on the training data was used. Also, for clarify the effect of each of the main features considered to predict mortality from acute respiratory syndrome of Covid-19, principal component analysis was used in Spss22. This nonparametric analysis is a method for selecting features with the aim of maximizing the prediction accuracy of regression and classification algorithms. Finally, according to the desired problem, five machine-learning algorithms such as random forest (RF), gradient boosting tree (GBT), support vector machine (SVM), and regression logistic (RL) were used in the Python 21 software using the relevant library functions.

2.3 | Modeling for Covid-19 mortality prediction

In the design of the SVM, we chose kernel function, “rbf,” to improve the efficiency of the network search, from the heuristic search

TABLE 1 Effective variables in Covid-19 mortality.

	Features name	Description	Type	Values
Demographic	Age	Age at diagnose	Demographic	<100 years
	Sex	Age of menopause	Demographic	38–65 years
	BMI	Body mass	Demographic	Underweight (below 18.5) = 0, normal (18.5–24.9) = 1, overweight (25.0–29.9) = 2, obese (30.0 and above) = 3
Patient risk factor history	Pulmunary disease	Diseases that cause airflow blockage	Demographic	Yes = 1 no = 0
	Drug smoke hist	History of smoking	Demographic	Yes = 1 no = 0
	Blood pressure	The force that moves blood through the circulatory system	Demographic	Yes = 1 no = 0
	Diabetes	Chronic disease	Demographic	Yes = 1 no = 0
	Cardiovascular disease	Conditions affecting the heart or blood vessels	Demographic	Yes = 1 no = 0
	Stroke	Poor blood flow to the brain that causes cell mortality	Demographic	Yes = 1 no = 0
	Chronic renal disease	A condition in which the kidneys are damaged	Laboratory	Yes = 1 no = 0
	Pregnancy		Laboratory	Yes = 1 no = 0
	Cancer	Personal other cancer	Laboratory	Yes = 1 no = 0
	COPD	A group of diseases that cause airflow blockage	Laboratory	Yes = 1 no = 0
	HIV	Human immunodeficiency virus	Laboratory	Yes = 1 no = 0
	Immunodeficiency	Absence of elements of the immune system	Laboratory	Yes = 1 no = 0
	Chemotherapy	History of chemotherapy	Laboratory	Yes = 1 no = 0
	Hyporthyroidism	Occurs when your thyroid gland produces too much of the hormone thyroxine	Laboratory	Yes = 1 no = 0
	AO_2 saturation		Clinical symptoms	Yes = 1 no = 0
	Muscle pain		Clinical symptoms	Yes = 1 no = 0
HRCT scan result for lung invasive		Screening	A: ground glass opacity, B: air bubble cyst, C: vessel related nodule, D: pleural effusion	0 = no findings; 1 < 25%; 2 = 25%–50%; 3 = 50%–75%; 4 > 75%
Laboratory features	PaO_2	Blood gas	Laboratory	75–100 mmHg
	$PaCO_2$	Blood gas	Laboratory	(35–45 mmHg)
	PaO_2:FIO_2	Blood gas	Laboratory	(35–45 mmHg)
	HCO_3	Blood gas	Laboratory	(22–26 meq/L)
	PH	Blood gas	Laboratory	(7.35–7.45)
	ICU duration			days
	Lymphocyte		Laboratory	20–40
	Creatinine		Laboratory	0.7–1.3 mg/dL
	FBS		Laboratory	60–200
	Hb1c		Laboratory	4%–10.5%

(Continues)

TABLE 1 (Continued)

	Features name	Description	Type	Values
	Na		Laboratory	136–146
	K		Laboratory	3.5–5
	LDH		Laboratory	Yes = 1 no = 0
	CRP		Laboratory	Yes = 1 no = 0
	Vitamin D		Laboratory	Yes = 1 no = 0
	D-dimer		Laboratory	≤1000, 1001–2500, >2500
	Pro calcitonin		Laboratory	Yes = 1 no = 0
	Ferritin		Laboratory	Yes = 1 no = 0
Class	Hospitalize status			Mortality = 1 discharge = 0

Abbreviations: ALT, alanine aminotransferase; AST, aspartate aminotransferase; BMI, body mass index; BUN, blood urea nitrogen; CA, calcium; COPD, chronic obstructive pulmonary disease; CRP, C-reactive protein; d-dimer, a fibrin degradation product, a small protein fragment present in the blood after a blood clot is degraded by fibrinolysis; ESR, erythrocyte sedimentation rate; FBS, fasting blood sugar; FIO_2, fractional inspired oxygen; Hb1c, hemoglobin A1c; HCO_3, carbonic acid then dissociates to form bicarbonate and hydrogen ions; HRCT, high-resolution computed tomography; ICU, intensive care unit; K, potassium; LDH, lactic dehydrogenase; Mg, magnesium; Na, sodium; $PaCO_2$, partial pressure of carbon dioxide; PaO_2, partial pressure of oxygen; Pro BNP, N-terminal pro B-type natriuretic peptide; PT, prothrombin time; PTT, partial thromboplastin time; RBC, red blood cells.

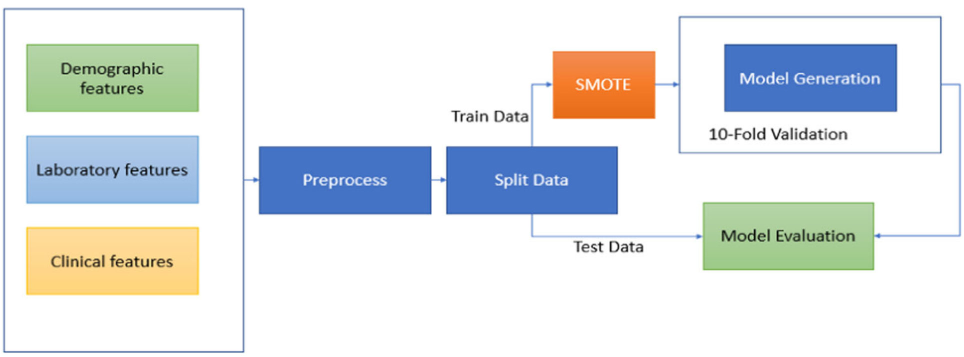

FIGURE 1 Design model diagram and analysis.

methods called the two-point center vertical method and the multipoint barycenter method, which can be used in GA, PSO, and Colony algorithm. In GBT, we considered (min_samples_split) minimum number of samples (or observations) equal to 10, (max_depth) or maximum depth of a tree equal to 5, maximum number of nodes (max_leaf_nodes) equal to 25. In RF, we considered the function of measuring the quality of a division as "gini," "entropy," the maximum depth of the tree equal to 5, the minimum number of samples required to divide an internal node equal to 2. In RL or reinforcement learning, we set the training rate (1–0.5) and the overall reward function includes immediate and delayed rewards. Its parameter is $0 < \gamma < 1$.

2.4 | Model evaluation

Using the K-Fold Cross Validation method ($K = 10$), nine parts of the data were selected as the training set and one part of the data as the test set. To assess the performance of prediction models, accuracy, sensitivity, specificity and ROC area under curve were calculated.

Figure 1 illustrates that the process of model development and evaluation. Figure 2, shows a histogram with the variable weight categorized. In addition, the final results of modeling analysis are shown in Table 2 and the scatter plots of partial thromboplastin time (PTT), D-dimer, and vitamin D variables are presented in Figure 4A–C.

Accuracy: (TP + TN)/(TP + TN + FP + FN)
Sensitivity: TP/(TP + FP)
Specificity: TN/(TN + FP)

As can be seen in the figure, age has a direct relationship with the progression of Covid-19, and having a high body mass can aggravate the symptoms of the disease. Furthermore, diabetes, heart diseases, immune system deficiency, and high blood concentration have a direct effect on the development of the symptoms of the disease.

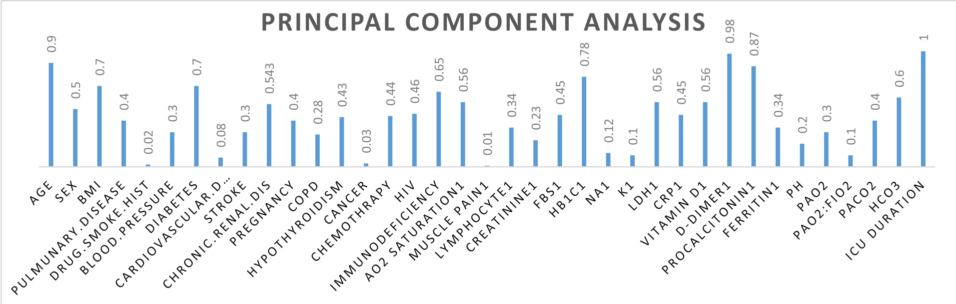

FIGURE 2 Weighting diagram used in modeling.

TABLE 2 The results of the models analysis.

Models	AUC	Sensitivity (%)	Specificity (%)	Accuracy (%)
GBT	0.854 + /_0.024	80.47 + /_6.33	60.87 + /_4.61	64.32 + /_3.14
Random forest	0.834 + /_0.029	77.67 + /_7.33	69.47 + /_4.31	70.47 + /_2.33
RL	0.812 + /_0.032	85.47 + /_5.33	68.87 + /_5.61	60.45 + /_4.84
SVM	0.795 + /_0.042	71.47 + /_3.33	70.44 + /_4.33	70.47 + /_3.03

Abbreviations: GBT, gradient boosting tree; RL, regression logistic; SVM, support vector machine.

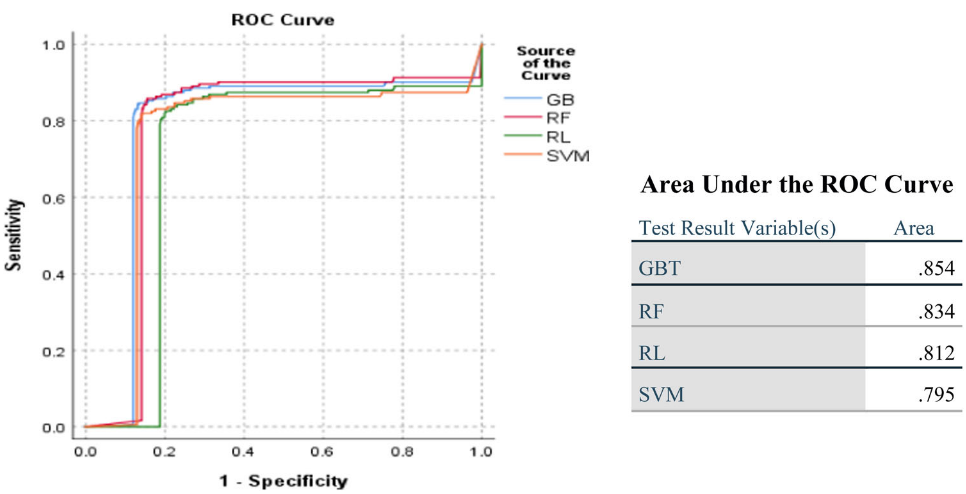

FIGURE 3 Performance diagram. GBT, gradient boosting tree; RF, random forest; RL, regression logistic; SVM, support vector machine.

Increasing the duration of stay in special care and using ventilator and artificial respiration can help the infection and flood of lung infection.

3 | RESULTS

Table 2 illustrates the performance of machine learning approaches with principal indicators such as sensitivity, accuracy, specificity, and AUC.

The comparison of the ROC area under curve of the four approaches used for predictive modeling is shown in Figure 3. The GBT model obtained a higher ROC area under curve than other models.

Figure 4 shows A: that most of the people with Covid-19 were in the normal range of the blood clotting test, B: most of the people over 60 years of age whose disease led to their mortality had blood concentration test results after contracting the disease and C: most of the people over 60 years of age whose disease led to their mortality had a vitamin D deficiency test result.

4 | DISCUSSION

In present study, four machine-learning approaches were used for predictive modeling in patients with Covid-19 considering 38 significant variables in patient's records. The results showed that

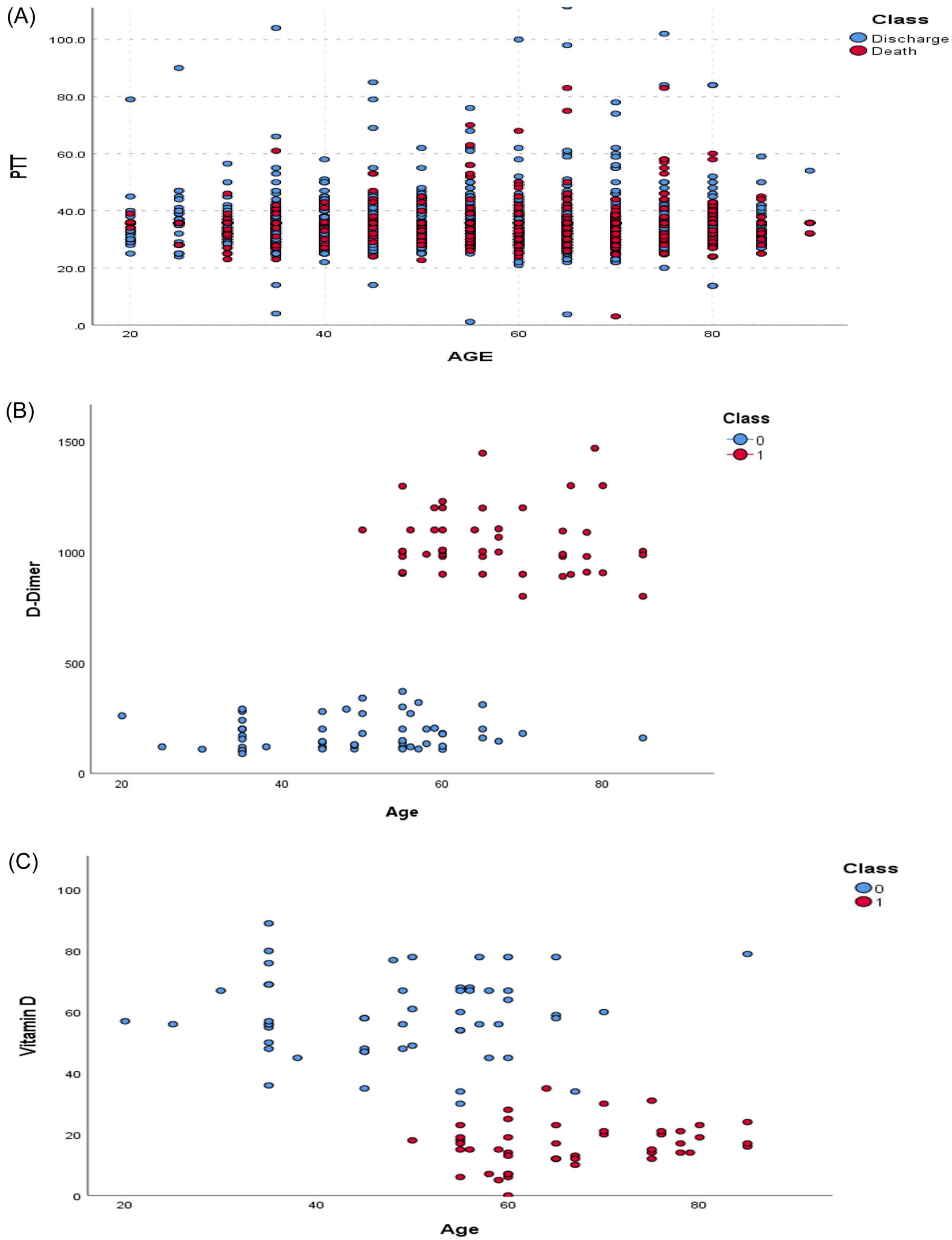

FIGURE 4 (A) Scattering of the amount of PTT test at different ages. (B) Scattering of D-dimer tests at different ages. (C) Scattering of vitamin D test levels at different ages. PTT, partial thromboplastin time.

GBT model obtained a higher prediction ability than other models, and its' area under the ROC curve obtained (0.854), respectively in this study.

In the analysis of the main variables used in the study, variables such as of length of hospitalization in the ICU, age, level of D-dimer test, autoimmune disease, level of vitamin D in patients had the highest weights. Measurement level of D-dimer (DD) and PTT were considered the significant indicators to be measured in patients with Covid-19 to diagnose thrombosis. In addition, comorbidity diseases such as diabetes, cancer, stroke, and pregnancy may increase the level of D-dimer and PTT tests in Covid-19 patients. Measuring D-dimer level and coagulation parameters in early stages of Covid-19 disease can be useful for controlling and managing of this disease, and in the current this contributing factors was considered as a significant indicator. Due to the importance of high levels of vitamin D in preventing hospitalization of patients, measuring this vitamin is very important.

Altschul et al.[21] studied the mortality of Covid-19 based on LR model using demographic, laboratory, and clinical data factors in a data set containing 4711 records, and the reported ROC curve was 0.79. In another study, Estiri et al.[22] used the GBT model on 830 records contain demographic and clinical features and reported ROC curve as 0.890. DAS et al.,[23] used five approaches such as GBT, NN, RF, LR, SVM; the reported accuracy for RF model was 98%. Laguna-Goya et al.[24] used the regression model on laboratory and demographic data to predict the risk of mortality with Covid-19, the reported ROC curve in their study was 0.947. Banoei et al.[25] developed the prediction model based on 108 clinical and paraclinical variables, the reported ROC curve for DT was 0.917. The findings of this study showed that the use of further variables can create higher performance models. In a study by Pourhomayoun,[26] the modeling on data of patients with Covid-19 was done using SVM, ANN, RF, DT, LR, and KNN, that the reported ROC curve was 0.897, which showed higher performance ability than other models. Yadaw et al.[27] investigated four technique such as LR, RF, SVM, and XGBT on 3841 records with 18 variables of patients with Covid-19. Models evaluating results shown that XGBT outperforms other models by obtaining an AUC of 0.915. Another similar work was done by Karthikeyan et al.,[28] used neural networks, logistic regression, XGBoost, RFs, SVM, and decision trees on 2779 records with 74 variables for prediction covid-19 mortality. The reported ROC curve in their study for LR was 99.26%. In a similar study by Rahman et al.,[29] proposed different machine learning classifiers: RF, SVM, KNN, XGBoost, and LR on 375 records patients with Covid-19 with 20 features. Logistic regression was the best performing machine learning classifier, by achieving an accuracy of 92.72%, sensitivity of 91.33%, and specificity of 78.26%. Gao et al.,[30] study, introduced four machine learning technique such as LR, SVM, KNN, RF, GBDT, and NN on database contain 2525 records, with 53 features including epidemiological, demographic, clinical, laboratory, radiological, RL model gives better results, that AUC reported 0.96, respectively. in a similar study by Chowdhury et al.[31] to predict Covid-19 mortality rate used RL and multitree XGBoost on a database consisting 2389 cases and 76 clinical features, the reported AUC for RL was 0.92. For prediction of mortality in Covid-19, Hu et al.[32] which approached the 10-fold evaluation method for RL, RF, and bagged FDA, the findings of this study showed that the RF led to an improved performance for models, the reported ROC curve in their study was 0.922. In a study by Tabatabaie et al.[33] analyzed six machine learning technique such as SVM; RF; DT, KNN; NB AdaBoost, on 520 records patients with covid-19 with 22 demographic and clinical features Bayes and neural network combination had slightly better performance with an AUC of 0.86, that it shows a respectable achievement in this study. Halasz et al.[34] predicted mortality with Covid-19 using a database containing 852 records with 14 features, including demographic risk factors and clinical data; their findings suggested that the Naïve Bayes with different factors showed improved performance (AUC = 0.808) compared to RF model. In the study by Booth et al.[35] proposed SVM technique such as on a database with 26 demographic and laboratory features. The findings indicated SVM model achieved 91% sensitivity and 91% specificity (AUC: 0.93) Subudhi et al.[36] performed 18 machine learning algorithms belonging to nine broad categories, namely ensemble, Gaussian process, linear, naïve Bayes, nearest neighbor, SVM, tree-based, discriminant analysis, and neural network models on a database with 3597 records and 50 variables to predict the mortality of patients with Covid-19 admitted to the ICU and also for ensemble-based methods have moderately better performance than other machine learning algorithms. The reported AUC was 0.99. The findings of the current study suggest that modeling with a variety of related risk factors from different sources could improve the performance of models in the mortality of patients with Covid-19 prediction.

5 | CONCLUSIONS

Machine learning has been applied in different field over the last years, and in medicine it has the potential to facilitate prediction of disease or could support clinicians in their decisions. With respect to Covid-19, machine learning techniques could be used in predicting the mortality of rate of this disease and taking necessary actions for better service delivery to patients.

6 | LIMITATIONS

There were limitations with the current study, as the modeling was done on records of only one database. In addition, we had no access to genetic data and such accessibility could help to draw a better picture of the performance of the techniques used. However, the use of different machine learning approaches and considering demographic, laboratory, and HRCT scan features, could provide a more robust basis for predicting Covid-19 mortality rate.

AUTHOR CONTRIBUTIONS

Hassan Emami: Conceptualization; formal analysis; investigation; methodology; project administration. **Reza Rabiei**: Conceptualization; funding acquisition; methodology; supervision. **Solmaz Sohrabei**: Methodology; resources; software; supervision; validation; visualization; writing—original draft; writing—review & editing. **Alireza Atashi**: Data curation.

ACKNOWLEDGMENTS

The research team are grateful to hospitals affiliated to TUMS for facilitating access to databases used in this study.

CONFLICT OF INTEREST STATEMENT

The authors declare no conflict of interest.

DATA AVAILABILITY STATEMENT

The authors confirm that the data supporting the findings of this study are available within the article and/or its supporting information.

ETHICS STATEMENT

This study was approved by the ethics committee of Tehran University of Medical Sciences (IR.TUMS.REC.140106050).

TRANSPARENCY DECLARATION

The lead author (Reza Rabiei and Solmaz Sohrabei) affirms that this manuscript is an honest, accurate, and transparent account of the study being reported; that no important aspects of the study have been omitted; and that any discrepancies from the study as planned (and, if relevant, registered) have been explained.

ORCID

Reza Rabiei http://orcid.org/0000-0003-0771-7306
Solmaz Sohrabei http://orcid.org/0000-0002-0736-5278

REFERENCES

1. Mahdavi M, Choubdar H, Zabeh E, et al. A machine learning based exploration of covid-19 mortality risk. *PLoS One*. 2021;16(7):e0252384.
2. Kuno T, Sahashi Y, Kawahito S, Takahashi M, Iwagami M, Egorova NN. Prediction of in-hospital mortality with machine learning for COVID-19 patients treated with steroid and remdesivir. *J Med Virol*. 2022;94(3):958-964.
3. Vaid A, Somani S, Russak AJ, et al. Machine learning to predict mortality and critical events in a cohort of patients with COVID-19 in New York city: model development and validation. *J Med Internet Res*. 2020;22(11):e24018.
4. Barda N, Riesel D, Akriv A, et al. Developing a covid-19 mortality risk prediction model when individual-level data are not available. *Nat Commun*. 2020;11(1):4439.
5. Cheng J, Sollee J, Hsieh C, et al. COVID-19 mortality prediction in the intensive care unit with deep learning based on longitudinal chest X-rays and clinical data. *Eur Radiol*. 2022;32(7):4446-4456.
6. Stachel A, Daniel K, Ding D, Francois F, Phillips M, Lighter J. Development and validation of a machine learning model to predict mortality risk in patients with COVID-19. *BMJ Health Care Informatics*. 2021;28(1):1-10.
7. Alizadehsani R, Alizadeh Sani Z, Behjati M, et al. Risk factors prediction, clinical outcomes, and mortality in COVID-19 patients. *J Med Virol*. 2021;93(4):2307-2320.
8. Das AK, Mishra S, Saraswathy Gopalan S. Predicting covid-19 community mortality risk using machine learning and development of an online prognostic tool. *PeerJ*. 2020;8:e10083.
9. El-Solh AA, Lawson Y, Carter M, El-Solh DA, Mergenhagen KA. Comparison of in-hospital mortality risk prediction models from COVID-19. *PLoS One*. 2020;15(12):e0244629.
10. Elhazmi A, Al-Omari A, Sallam H, et al. Machine learning decision tree algorithm role for predicting mortality in critically ill adult COVID-19 patients admitted to the ICU. *J Infection Public Health*. 2022;15(7):826-834.
11. Albitar O, Ballouze R, Ooi JP, Sheikh Ghadzi SM. Risk factors for mortality among covid-19 patients. *Diabetes Res Clin Pract*. 2020;166(108293):108293.
12. Cao Y, Li L, Feng Z, et al. Comparative genetic analysis of the novel coronavirus (2019-nCoV/SARS-CoV-2) receptor ACE2 in different populations. *Cell Discov*. 2020;6(1):1-4.
13. Asselta R, Paraboschi EM, Mantovani A, Duga S. ACE2 and TMPRSS2 variants and expression as candidates to sex and country differences in covid-19 severity in Italy. *Aging*. 2020;12(11):10087-10098.
14. Dessie ZG, Zewotir T. Mortality-related risk factors of COVID-19: a systematic review and meta-analysis of 42 studies and 423,117 patients. *BMC Infect Dis*. 2021;21(1):855.
15. Williamson EJ, Walker AJ, Bhaskaran K, et al. Factors associated with COVID-19-related death using OpenSAFELY. *Nature*. 2020;584(7821):430-436.
16. Schultze A, Walker AJ, MacKenna B, et al. Risk of COVID-19-related death among patients with chronic obstructive pulmonary disease or asthma prescribed inhaled corticosteroids: an observational cohort study using the OpenSAFELY platform. *Lancet Respiratory Med*. 2020;8(11):1106-1120.
17. Arvind V, Kim JS, Cho BH, Geng E, Cho SK. Development of a machine learning algorithm to predict intubation among hospitalized patients with COVID-19. *J Crit Care*. 2021;62:25-30.
18. Hashim MJ, Alsuwaidi AR, Khan G. Population risk factors for covid-19 mortality in 93 countries. *J Epidemiol Global Health*. 2020;10(3):204-208.
19. Noor FM, Islam MM. Prevalence and associated risk factors of mortality among covid-19 patients: a meta-analysis. *J Community Health*. 2020;45(6):1270-1282.
20. Freuer D, Linseisen J, Meisinger C. Impact of body composition on covid-19 susceptibility and severity: a two-sample multivariable Mendelian randomization study. *Metabolism*. 2021;118:154732.
21. Altschul DJ, Unda SR, Benton J, et al. A novel severity score to predict inpatient mortality in covid-19 patients. *Sci Rep*. 2020;10(1):16726.
22. Estiri H, Strasser ZH, Klann JG, Naseri P, Wagholikar KB, Murphy SN. Predicting covid-19 mortality with electronic medical records. *NPJ Digital Med*. 2021;4(1):15.
23. Malki Z, Atlam ES, Hassanien AE, Dagnew G, Elhosseini MA, Gad I. Association between weather data and COVID-19 pandemic predicting mortality rate: machine learning approaches. *Chaos Solitons Fractals*. 2020;138(110137):110137.
24. Laguna-Goya R, Utrero-Rico A, Talayero P, et al. IL-6–based mortality risk model for hospitalized patients with COVID-19. *J Allergy Clin Immunol*. 2020;146(4):799-807.
25. Banoei MM, Dinparastisaleh R, Zadeh AV, Mirsaeidi M. Machine-learning-based covid-19 mortality prediction model and identification of patients at low and high risk of dying. *Crit Care*. 2021;25(1):328.
26. 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.
27. Yadaw AS, Li Y, Bose S, Iyengar R, Bunyavanich S, Pandey G. Clinical features of covid-19 mortality: development and validation of a clinical prediction model. *Lancet Digital Health*. 2020;2(10):e516-e525.
28. Karthikeyan A, Garg A, Vinod PK, Priyakumar UD. Machine learning based clinical decision support system for early covid-19 mortality prediction. *Front Public Health*. 2021;9:1-13.
29. Rahman T, Al-Ishaq FA, Al-Mohannadi FS, et al. Mortality prediction utilizing blood biomarkers to predict the severity of covid-19 using machine learning technique. *Diagnostics*. 2021;11(9):1582.
30. Gao Y, Cai GY, Fang W, et al. Machine learning based early warning system enables accurate mortality risk prediction for COVID-19. *Nat Commun*. 2020;11(1):5033.

31. Chowdhury M, Rahman T, Khandakar A, et al. An early warning tool for predicting mortality risk of covid-19 patients using machine learning. *Cognitive Computation*. 2021;21:1-16.
32. Hu C, Liu Z, Jiang Y, et al. Early prediction of mortality risk among patients with severe COVID-19, using machine learning. *Int J Epidemiol*. 2021;49(6):1918-1929.
33. Tabatabaie M, Sarrami AH, Didehdar M, Tasorian B, Shafaat O, Sotoudeh H. Accuracy of machine learning models to predict mortality in COVID-19 infection using the clinical and laboratory data at the time of admission. *Cureus*. 2021;13(10):1-6.
34. Halasz G, Sperti M, Villani M, et al. A machine learning approach for mortality prediction in covid-19 pneumonia: development and evaluation of the Piacenza score. *J Med Internet Res*. 2021;23(5):e29058.
35. Booth AL, Abels E, McCaffrey P. Development of a prognostic model for mortality in covid-19 infection using machine learning. *Mod Pathol*. 2021;34(3):522-531.
36. Subudhi S, Verma A, Patel AB, et al. Comparing machine learning algorithms for predicting ICU admission and mortality in COVID-19. *NPJ Digital Med*. 2021;4(1):87.

How to cite this article: Emami H, Rabiei R, Sohrabei S, Atashi A. Predicting the mortality of patients with Covid-19: a machine learning approach. *Health Sci Rep*. 2023;6:e1162. doi:10.1002/hsr2.1162