Research Paper

Risk Score Model for Predicting COVID-19 Progression in Iranian Patients: Development and Validation Study

Fatemeh Mohammadzadeh[1], Ali Delshad Noughabi[2], Sina Sabeti Bilondi[3,4], Mitra Tavakolizadeh[4], Jafar Hajavi[5], Hosein Aalami[6], Mohsen Sahebanmaleki[7*]

1. Department of Epidemiology and Biostatistics, Social Development and Health Promotion Research Center, School of Health, Gonabad University of Medical Sciences, Gonabad, Iran.
2. Social Development and Health Promotion Research Center, Gonabad University of Medical Sciences, Gonabad, Iran.
3. Department of Nursing, Faculty of Nursing and Midwifery, Gonabad Branch, Islamic Azad University, Gonabad, Iran.
4. Unit of Clinical Research Development, Bohlool Hospital, Gonabad University of Medical Sciences, Gonabad, Iran.
5. Department of Microbiology, Infectious Disease Research Center, School of Medicine, Gonabad University of Medical Sciences, Gonabad, Iran.
6. Unit of Clinical Research Development, Gonabad University of Medical Sciences, Gonabad, Iran.
7. Health Promotion Research Center, Bohlool Hospital, Gonabad University of Medical Sciences, Gonabad, Iran.

Use your device to scan and read the article online

Citation Mohammadzadeh F, Delshad Noughabi A, Sabeti Bilondi S, Tavakolizadeh M, Hajavi J, Aalami H, et al. Risk Score Model for Predicting COVID-19 Progression in Iranian Patients: Development and Validation Study. Journal of Research & Health. 2024; 14(3):217-230. http://dx.doi.org/10.32598/JRH.14.3.646.1

http://dx.doi.org/10.32598/JRH.14.3.646.1

ABSTRACT

Background: The recent novel coronavirus disease 2019 (COVID-19) pandemic has underlined the importance of risk score models in public health emergencies. This study aimed to develop a risk prediction score to identify high-risk hospitalized patients for disease progression on admission.

Methods: This prospective cohort study included 171 COVID-19 patients, identified through the reverse transcription polymerase chain reaction test, admitted to Bohlool Hospital in Gonabad City, Iran, between April 4 and June 5, 2021. The patients' demographic, clinical, and laboratory data were collected upon admission, and clinical outcomes were monitored until the end of the study. The discovery dataset (80% of the data) was used to develop the risk score model based on clinical and laboratory features and patient characteristics to predict COVID-19 progression. An additive risk score model was developed based on the regression coefficients of the significant variables in a multiple logistic regression model. The performance of the risk score model was evaluated on the validation dataset (20% of the data) using the receiver operating characteristic (ROC) curve. Statistical analyses were performed with SPSS software, version 21.

Results: The Mean±SD for age of participants was 59.54±20.52 years, and 48.6% were male. Most patients (82.5%) fully recovered or showed improvement, while 5.2% experienced disease progression and 12.3% died. Three variables, interleukin-6, neutrophil-to-lymphocyte ratio, and lung involvement, were found to be significant in predicting risk, with a good discriminatory ability, having an area under the ROC curve of 0.970 (95% CI, 0.935%, 1.00%) in the discovery set and 0.973 (95% CI, 0.923%, 1.00%) in the validation set.

Conclusion: The developed risk score model in this study can be used as a clinical diagnostic tool to identify COVID-19 patients at higher risk of disease progression and aid in informed decision-making and resource utilization in similar situations, such as respiratory disease outbreaks in the post-corona era.

Keywords: Coronavirus, COVID-19, Risk score, Prediction, Disease progression, Iran

Article info:
Received: 31 Jul 2023
Accepted: 07 Oct 2023
Publish: 01 May 2024

*** Corresponding Author:**
Mohsen Sahebanmaleki, MD.
Address: *Health Promotion Research Center, Bohlool Hospital, Gonabad University of Medical Sciences, Gonabad, Iran.*
Phone: *+98 (51) 57229025*
E-mail: *dr.saheban@yahoo.com*

Introduction

In late 2018, the novel coronavirus disease 2019 (COVID-19) emerged in China and spread rapidly worldwide, becoming a crucial global public health emergency with an unprecedented strain on health and economic systems [1, 2]. The prevalence of new variants of coronavirus like Delta and Omicron posed more challenges in managing the pandemic due to their high transmissibility, ability to evade detection, and resistance to vaccines [3]. These variants placed additional pressure on limited health resources and led to more hospitalizations and deaths [4]. As of 26 July 2023, a total of 773511195 confirmed cases of COVID-19, including 7023127 deaths, were reported to the World Health Organization (WHO) [5].

A risk prediction model is a mathematical equation that makes use of patient data to assess the likelihood of encountering a healthcare outcome. There are various statistical techniques available for creating a risk prediction model. Some common models include the logistic regression model, Cox proportional hazards model, and classification trees [6, 7]. Logistic regression is a traditional method widely used for binary outcomes and is renowned for its strong explanatory ability. It frequently utilizes a variable selection approach, such as the backward selection method, to identify the strongest predictors. The logistic models were frequently converted into straightforward additive risk scores by assigning weights to the predictors according to the log odds ratios derived from the model to simplify their usability [6]. The risk score models are widely employed in medical practice to forecast the development of diseases, assess treatment effectiveness, predict patient prognosis, or identify individuals at a higher risk of disease progression and mortality [4]. By providing risk estimates, these models play a crucial role in guiding clinical decisions regarding the limited resource and capacity allocation to patients most likely to benefit from early intervention. They aid in preventing disease progression, reducing the risk of complications, and alleviating the burden on healthcare systems [4]. The recent COVID-19 pandemic has underscored the significance of risk prediction models in public health emergencies.

So far, several models have been proposed based on integrating demographic, clinical, and radiological features for early prediction and identification of patients at risk for severe pneumonia, intubation, intensive care unit (ICU) transfer, and patient death [3, 4, 8, 9].

However, these models have not generally been used in clinical practice. They are prone to bias [4] due to several reasons, such as a limited selection of samples, the retrospective nature of the study design, and unclear details regarding the development and validation of the model. However, it is essential to note that with new variants of SARS-CoV-2 with distinct characteristics, these models need to be updated accordingly. Iran has been significantly impacted by the COVID-19 pandemic. Nevertheless, limited clinical experiences and research have been conducted on the Iranian population. Only a few studies are available that have retrospectively examined the development of risk score prediction models for COVID-19 in Iranian patients [10-12], which diminishes the reliability and validity of these studies. Therefore, we undertook a prospective study to address this gap and develop a more reliable and validated risk prediction score model to identify Iranian COVID-19 patients at a higher risk of disease progression during hospitalization.

Methods

Study design, participants, and data collection

This cohort study was conducted at Bohlool Hospital in Gonabad City, Iran, from April 4 to June 5, 2021. The study included COVID-19 patients admitted to the hospital who met the eligibility criteria. The patients' demographic, clinical, and laboratory information was collected upon admission. The patients were followed over time, and their clinical outcomes were recorded until the end of the study. The inclusion criteria comprised a definitive diagnosis of COVID-19 by reverse transcription-polymerase chain reaction (RT-PCR) test and a minimum age of 18 years. All patients were treated according to the physician's diagnosis. White blood cells (WBC), neutrophil to lymphocyte ratio (NLR), interleukin-6 (IL-6), C-reactive protein (CRP), lactate dehydrogenase (LDH), and ferritin tests were performed as a supplement for all patients. Other data were extracted using hospital information system (HIS) resources. Patients with incomplete records were excluded from the study.

Potential predictive variables

Potential predictor variables were as follows: Age, sex, pregnancy, use of tobacco, use of opium, history of COVID-19, inpatient department, partial pressure of oxygen (PaO_2), temperature, computed tomographic (CT) imaging score, signs and symptoms at admission (including fever, cough, muscular pain, respiratory distress, loss of consciousness, decreased sense of smell, reduced sense of taste, seizure, abdominal pain, nausea,

vomiting, diarrhea, anorexia, headache, vertigo, paresthesia, paraplegia, chest pain, skin lesion(, comorbidities (liver diseases, diabetes, hematologic diseases, human immunodeficiency virus/acquired immune deficiency syndrome [HIV/AIDS], autoimmune diseases, heart diseases, kidney diseases, asthma, chronic lung diseases, nervous diseases, hypertension, and other), laboratory values (WBC, NLR, CRP, IL-6, LDH, and ferritin), and clinical management and pharmacological treatment (tracheal intubation, O_2 therapy, and drugs).

CT imaging scoring

An experienced radiologist classified CT images into four categories based on the extent of pulmonary parenchymal involvement: Mild, moderate, severe, and critical. The mild stage included CT images with several ground-glass opacities (GGOs) and less than 25% lung tissue involvement. The moderate stage included CT images with GGOs and between 25% and 50% of lung tissue involvement. The severe stage showed both GGOs and areas of consolidation, with lung tissue involvement ranging from 50% to 75%. The critical stage included diffuse GGOs, consolidation, and reticular changes in the lungs, with lung tissue involvement exceeding 75% [13].

Outcomes

Patients were divided into two groups based on the outcomes: 1) Without disease progression, patients who fully recovered and discharged or showed stable symptomatic improvement, and 2) With disease progression, patients who had progressed to severe illness and stayed in the intensive care unit or died.

Ethical considerations

The study was approved by the Ethics Committee of Gonabad University of Medical Sciences. The ethical principles of human medical research (Helsinki Declaration) were observed in all study phases. All patients' data were extracted confidentially and encrypted from the hospital information system (HIS). The patient's diagnosis process was carried out under the guidelines published by the WHO and the Iranian Ministry of Health and Medical Education.

Statistical analysis

Descriptive statistics and bivariate analysis

The Kolmogorov-Smirnov test was used to examine the normal distribution for quantitative variables, including age, PaO_2, temperature, duration of hospitalization, and laboratory values. The age distribution was reported as the Mean±SD. The two groups with and without disease progression were compared using the independent t-test. The median (25^{th} percentile, 75^{th} percentile) and the Mann-Whitney test were used to describe and compare other quantitative variables between the two groups, respectively. Describing and comparing qualitative variables were carried out using number (percentage) and the chi-square test, respectively. A two-sided $P<0.05$ was considered significant.

Development and validation of risk score prediction model

After randomly dividing data into two parts (the discovery dataset with 80% of data and the validation dataset with 20%), the prediction model was developed using a logistic regression model on the discovery dataset. The variables with a $P<0.2$ in the simple logistic regression model were entered into a multiple logistic regression model. We considered a backward removal method with $P<0.05$ for entering variables and $P<0.1$ for removing variables into and from the multiple logistic regression model. Then, the coefficients obtained from the model were converted to an integer risk score. The highest sensitivity and specificity values were used to determine the optimum cut-off point for the risk score model [14]. Model calibration was assessed using the Hosmer-Lemeshow test with a $P>0.05$, indicating acceptable goodness of fit to the data [15]. The area under the receiver operator curve (AUC), with a minimum value of 0.70 as a desirable discrimination ability [14], was used to evaluate the discrimination ability of the risk score model in both discovery and validation datasets. The performance of the risk score prediction model was compared with each of the predictors used in the model, namely IL-6, NLR, and lung involvement. Additionally, the performance of the risk score was compared to two other laboratory variables, CRP and LDH, which were statistically significant in the bivariate analyses. The significance level was considered 0.05. All data analysis was performed using SPSS software, version 21.

Results

Individual and clinical characteristics of patients

Data were analyzed from 171 patients with COVID-19, of which 149(87.13%) were hospitalized in the isolation ward and 22(12.86%) in the ICU. Of 171 patients, 141(82.5%) were recovered or showed symptomatic improvement. Nine patients (5.2%) had progressed to severe disease, while 21(12.3%) died. The individual and clinical characteristics of the patients are shown in Table 1.

Table 1. Demographic and clinical characteristics of hospitalized COVID-19 patients

Characteristics			Mean±SD/No.(%)/Median		P
			Disease Progression		
			Yes	No	
Age (y)			71.90±14.3	59.33±17.7	<0.001†#
Sex		Man	18(60.0)	68(48.2)	0.315†
		Woman	12(40.0)	73(51.8)	
Pregnancy		Yes	1(3.3)	0(0.0)	0.175†
		No	29.0(96.7)	141(100.0)	
Use tobacco		Yes	1(3.3)	1(0.7)	0.321†
		No	26(96.7)	140(99.3)	
Use opium		Yes	2(6.7)	4(2.8)	0.592†
		No	28(93.3)	137(97.2)	
Inpatient department		Isolation room	15(50.0)	134(95.0)	<0.001†#
		ICU	15(50.0)	7(5.0)	
Radiological finding	History of COVID-19	Yes	2(6.7)	13(9.2)	0.745†
		No	28(93.3)	128(90.8)	
	Lung involvement	Mild	6(20)	38(27)	<0.001†#
		Moderate	7(23.3)	50(35.5)	
		Severe	14(46.7)	8(5.7)	
		Critical	2(6.7)	1(0.7)	
Signs and symptoms at admission	Fever	Yes	6(20.0)	61(43.3)	0.018†#
		No	24(80.0)	80(56.7)	
	Cough	Yes	9(30.0)	54(38.3)	0.392†
		No	21(70.0)	87(61.7)	
	Muscular pain	Yes	3(10.0)	17(12.1)	0.777†
		No	27(90.0)	124(87.9)	
	Respiratory distress	Yes	24(80.0)	80(56.7)	0.018†#
		No	6(20.0)	61(43.3)	
	Loss of consciousness	Yes	3(10.0)	6(4.3)	0.363†
		No	27(90.0)	135(95.7)	
	Decreased sense of smell	Yes	0(0.0)	4(2.8)	0.601†
		No	30(100)	137(97.2)	

Characteristics			Mean±SD/No.(%)/Median		P
			Disease Progression		
			Yes	No	
Signs and symptoms at admission	Decreased sense of taste	Yes	0(0.0)	4(2.8)	0.601†
		No	30(100)	137(97.2)	
	Seizure	Yes	1(0.70)	4(2.8)	1.000†
		No	140(99.30)	137(97.2)	
	Abdominal pain	Yes	1(3.3)	4(2.8)	1.000†
		No	29(96.7)	137(97.2)	
	Nausea	Yes	0(0.0)	17(12.1)	0.084†
		No	30(100)	124(87.9)	
	Vomiting	Yes	0(0.0)	5(3.5)	0.588†
		No	30(100)	136(96.5)	
	Diarrhea	Yes	0(0.0)	4(2.8)	0.601†
		No	30(100)	137(97.2)	
	Anorexia	Yes	2(6.7)	20(14.2)	0.374†
		No	28(93.3)	121(85.8)	
	Headache	Yes	0(0.0)	9(6.4)	0.220†
		No	30(100)	132(93.6)	
	Vertigo	Yes	0(0)	2(1.4)	1.000†
		No	30(100)	139(98.6)	
	Paresthesia	Yes	0(0.0)	0(0.0)	---
		No	30(100)	141(100)	
	Paraplegia	Yes	0(0.0)	0(0.0)	---
		No	30(100)	141(100)	
	Chest pain	Yes	1(3.3)	3(2.1)	1.000†
		No	29(96.7)	138(97.9)	
	Skin lesion	Yes	0(0.0)	0(0.0)	---
		No	30(100)	141(100)	
	PaO_2, Median (25^{th}, 75^{th})		87.00 (77.75, 91.00)	92.00 (89.00, 95.00)	<0.001*#
	Temperature, Median (25^{th}, 75^{th})		37.00 (36.50, 37.72)	37.00 (36.70, 37.80)	0.458*

Characteristics			Mean±SD/No.(%)/Median		P
			Disease Progression		
			Yes	No	
Comorbidities	Liver diseases	Yes	0(0.0)	1(0.7)	1.000†
		No	30(100)	140(99.3)	
	Diabetes	Yes	6(20.0)	19(13.5)	0.394†
		No	24(80.0)	122(86.5)	
	Hematologic diseases	Yes	0(0.0)	2(1.4)	1.000†
		No	30(100)	139(98.6)	
	HIV/AIDS	Yes	0(0.0)	0(0.0)	---
		No	30(100)	141(100)	
	Autoimmune diseases	Yes	0(0.0)	0(0.0)	---
		No	30(100)	141(100)	
	Heart diseases	Yes	5(16.7)	18(12.8)	0.570†
		No	25(83.3)	123(87.2)	
	Kidney diseases	Yes	0(0.0)	0(0.0)	---
		No	30(100)	141(100)	
	Asthma	Yes	3(10)	4(2.8)	0.104†
		No	27(90)	137(97.2)	
	Lung diseases	Yes	6(20)	5(3.5)	0.004†#
		No	24(80)	136(96.5)	
	Nervous diseases	Yes	1(3.3)	1(0.7)	0.321†
		No	29(96.7)	140(99.3)	
	Hypertension	Yes	16(53.3)	38(27.0)	0.005†#
		No	14(46.7)	103(73.0)	
	Other's diseases	Yes	5(16.67)	19(13.47)	1.000†
		No	25(83.3)	122(86.5)	
Laboratory values	WBC (mm^3) Median (25^{th}, 75^{th})		58.50 (44.75, 91.25)×10^2	53.00 (40.00, 72.50)×10^2	0.144*
	NLR Median (25^{th}, 75^{th})		4.93 (2.51, 12.08)	3.63 (2.37, 5.65)	0.032*#
	CRP (mg/L) Median (25^{th}, 75^{th})		47.36 (16.51, 101.43)	20.49 (5.66, 45.77	0.005*#
	IL-6 (pg/mL) Median (25^{th}, 75^{th})		23.90 (17.85, 33.50)	10.80 (9.10, 13.15)	<0.001*#
	LDH (U/L) Median (25^{th}, 75^{th})		381.50 (320.00, 586.00)	339.00 (265.00, 438.00)	0.027*#
	Ferritin (ng/L) Median (25^{th}, 75^{th})		237.30 (143.40, 374.22)	217.00 (101.80, 352.25)	0.542*

Characteristics			Mean±SD/No.(%)/Median		P
			Disease Progression		
			Yes	No	
Clinical management & pharmacological treatment	Tracheal intubation	Yes	19(63.3)	3(2.1)	<0.001†#
		No	11(36.7)	138(97.9)	
	O_2 therapy	Yes	6(20)	40(28.4)	0.348†
		No	24(80)	101(71.6)	
	Remdesivir	Yes	20(66.7)	117(83.0)	0.042†#
		No	10(33.3)	24(17.0)	
	Interferon	Yes	7(23.3)	45(31.9)	0.354†
		No	23(76.7)	96(68.1)	
	Dexamethasone	Yes	21(70.0)	101(71.6)	0.858†
		No	9(30.0)	40(28.4)	
	Methylprednisolone	Yes	14(46.7)	46(32.6)	0.143†
		No	16(53.3)	95(67.4)	
	Neurobion	Yes	2(6.7)	16(11.3)	0.448†
		No	28(93.3)	125(88.7)	
	Favipiravir	Yes	1(3.3)	2(1.4)	1.000†
		No	29(96.7)	139(98.6)	
	Hydrocortisone	Yes	4(13.3)	3(2.1)	0.019†#
		No	26(86.7)	138(97.9)	
	Atorvastatin	Yes	19(63.3)	87(61.7)	0.867†
		No	11(36.7)	54(38.3)	
	Aspirin	Yes	17(56.7)	77(54.6)	0.837†
		No	13(43.3)	64(45.4)	
	Chloroquine	Yes	4(13.3)	20(14.2)	1.000†
		No	26(86.7)	121(85.8)	
	Famotidine	Yes	18(60.0)	105(74.5)	0.109†
		No	12(40.0)	36(25.5)	
	Zinc	Yes	8(26.7)	45(31.9)	0.572†
		No	22(73.3)	96(68.1)	
	Heparin	Yes	27(90.0)	124(87.9)	0.777†
		No	3(10.0)	17(12.1)	
	Vitamin C	Yes	12(40.0)	54(38.3)	0.862†
		No	18(60.0)	87(61.7)	

Characteristics			Mean±SD/No.(%)/Median		P
			Disease Progression		
			Yes	No	
Outcomes	Duration of hospitalization (25th, 75th)		9.00 (3.00, 18.25)	5.00 (4.00, 7.00)	0.014*#
	Final outcome of COVID-19	Fully recover	0(0.0)	15(10.6)	
		Improved	0(0.0)	126(89.4)	
		Exacerbation	9(30.0)	0(0.0)	
		Death	21(70.0)	0(0.0)	

Abbreviations: PaO_2: Partial pressure of oxygen; CT: Computed tomographic; WBC: White blood cells; NLR: Neutrophil-to-lymphocyte ratio; CRP: C-reactive protein; IL-6: Interleukin-6; LDH: Lactate dehydrogenase.

ǂThe independent t-test, *The Mann-Whitney test, †The chi-square test, #P<0.05.

Bivariate analysis

The findings of bivariate analysis showed several distinguishing characteristics between patients whose COVID-19 progressed and those who were discharged or experienced symptomatic improvement. These characteristics included a higher age (P<0.001), lower median PaO_2 values upon admission (P<0.001), lower prevalence of fever (P=0.018), and higher prevalence of respiratory distress upon admission (P=0.018). Additionally, there was a higher prevalence of preexisting lung diseases (P=0.004) and hypertension (P=0.005), more severe lung involvement (P<0.001), elevated levels of NLR (P=0.032), CRP (P=0.005), IL-6 (P<0.001), and LDH (P=0.027), reduced administration of Remdesivir (P=0.042), and increased administration of hydrocortisone (P=0.019) among those whose condition progressed.

Development and validation of risk score prediction model

Sixteen variables, including age, sex, use of opium, inpatient department, fever, respiratory distress, loss of consciousness, lung involvement, hypertension, chronic lung disease, Remdesivir, hydrocortisone, IL-6, CRP, LDH, and NLR in the simple logistic regression model with P<0.2 were selected as potential predictor variables to enter the multiple logistic regression model. Three variables, including lung involvement, NLR, and IL-6, were significant in the multiple logistic regression model and were set in the risk score formula (Table 2). The risk score for disease progression was developed based on the regression coefficients of the multiple logistic regression model (Equations 1 and 2).

1. Risk score=(0.433×IL-6)+(0.205×NLR)+(0.072×Long involvement)

Table 2. Logistic regression model results of associated factors with adverse effects in patients hospitalized with COVID-19

Variables	OR (95% CI)		P	
	Unadjusted	Adjusted	Simple	Multiple
	Simple	Multiple		
IL-6	1.41 (1.25, 1.59)	1.54 (1.26, 1.89)	<0.001	<0.001
NLR	1.13 (1.05, 1.22)	1.23 (1.06, 1.43)	0.001	0.008
Lung involvement	1.05 (1.03, 1.08)	1.07(1.03, 1.12)	<0.001	0.001

Abbreviations: IL-6: Interleukin 6; NLR: Neutrophil-to-lymphocyte ratio; OR: Odds ratio; CI: Confidence interval.

The Equation 2 was obtained to calculate the probability of disease progression.

2. *Probability=Exp(Risk score)/1+Exp(Risk score)*

The Hosmer - Lemeshow test statistic indicated a well-fitting risk score model to the data (P=0.681).

The performance and validation of risk score

The AUC was 97.0 (95% CI, 93.5%, 1.00%) in the discovery dataset and 97.3 (95% CI, 92.3%, 100.0%) in the validation dataset, indicating a good performance in both discovery and validation datasets. The performance of combined variables of IL-6, NLR, and lung involvement as predictors of disease progression in the risk score prediction model was significantly higher than IL-6, NLR, lung involvement, CRP, and LDH alone (Figure 1). Table 3 listed the cut-off values of sensitivity, specificity, positive and negative predictive values of each of the mentioned variables, and AUC of the established prediction model.

Discussion

In this study, a risk score prediction model was established to aid in the identification of Iranian COVID-19 patients who are at a higher risk of disease progression at the time of hospitalization. The risk score model comprised three variables, IL-6, NLR, and lung involvement, and demonstrated an AUC of 97.0 (95% CI, 93.5%, 1.00%) in the discovery dataset and 97.3 (95% CI, 92.3%, 100.0%) in the validation dataset. These results indicate good performance in discovery and validation datasets, comparable to other models in the literature. A study has shown that predictive scores based on IL-6 and NLR, in combination with respiratory rate, SpO_2/FiO_2 ratio, and LDH, exhibit superior capability (AUC over 0.80) compared to other similar scores developed for the prediction of adverse outcomes in COVID-19, specifically in the prediction of invasive mechanical ventilation [16]. Another study has demonstrated that three models containing one of the parameters of NLR or IL-6, combined with three clinical parameters (age, sex, and SpO_2), exhibit promising discrimination in predicting adverse outcomes in COVID-19 patients [17]. In another study, IL-6 has been identified as an independent predictor of COVID-19 mortality in a six-point prediction score model [18]. Additionally, a study has verified the predictive role of chest CT severity score in the need for invasive mechanical ventilation and mortality in COVID-19 patients, with an AUC of 0.759 [19]. It is said that the increase in NLR values is related to the decrease in lymphocytes, which have a crucial role in battling SARS-infected cells. The direct effect of SARS-CoV-2 on lymphocytes causing cell death, increasing lactic acid and inhibiting lymphocyte proliferation, and directly invading and destroying the lymphatic organs are the main rea-

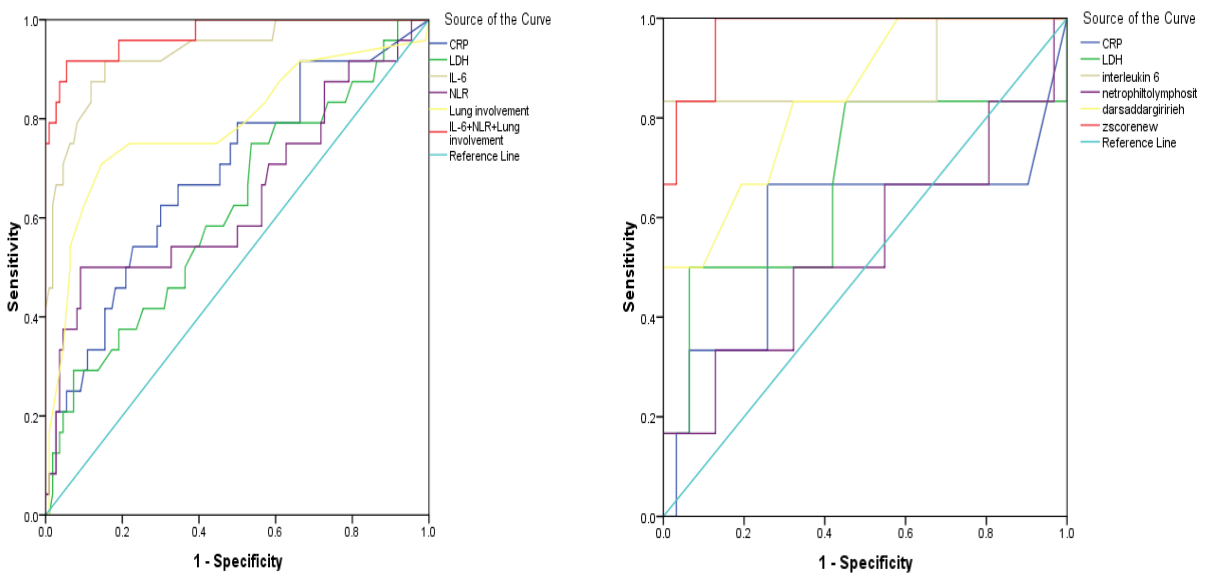

Figure 1. ROC curve of CRP, LDH, NLR, Lung involvement, and IL-6

Left: Discovery dataset, Right: Validation datasets

Abbreviations: ROC: Receiver operator curve; CRP: C-reactive protein; LDH: Lactate dehydrogenase; NLR: Neutrophil-to-lymphocyte ratio; IL-6: Interleukin-6.

Table 3. Sensitivity, specificity, NPV, PPV, and AUC of risk score and biomarkers in discovery and validation datasets

Marker	Discovery Dataset					
	CRP	LDH	NLR	Long Involvement	IL-6	Risk Score: Long Involvement+IL-6+NLR
Cut-off	33.22	356.50	4.54	42.5	14.45	-2.45

Marker		% (95% CI)					
		CRP	LDH	NLR	Long Involvement	IL-6	Risk Score: Long Involvement+IL-6+NLR
Discovery Dataset	Sensitivity	66.7 (44.7, 84.4)	58.3 (36.6, 78.0)	54.2 (32.8, 74.5)	70.8 (48.6, 86.6)	91.7 (73.0,99.0)	91.7 (73.0, 99.0)
	Specificity	65.5 (55.8, 74.3)	56.4 (47.0, 65.8)	60.0 (50.2, 69.2)	85.5 (77.1, 91.2)	84.5 (76.4, 90.7)	88.2 (80.6, 93.6)
	NPV	90.0 (83.4, 94.1)	86.1 (79.0, 91.1)	85.7 (79.1, 90.5)	93.1 (85.8, 96.9)	97.9 (92.4, 99.4)	98.0 (92.8, 99.5)
	PPV	29.6 (22.3, 38.1)	22.6 (16.4, 91.1)	22.8 (16.1, 31.3)	51.5 (33.6, 68.8)	56.4 (45.1, 67.1)	62.9 (50.0, 74.1)
	AUC	68.8 (56.7, 81.0)	61.8 (48.9, 74.7)	64.6 (50.6, 78.6)	78.7 (66.7, 90.7)	93.4 (87.7, 99.1)	97.0 (93.5, 100.0)
Validation Dataset	Sensitivity	66.7 (22.3, 95.7)	50.0 (11.8, 88.2)	50.0 (11.8, 88.2)	66.7 (24.1, 94.0)	83.3 (35.9, 99.6)	100.0 (54.1, 100.0)
	Specificity	64.5 (45.4, 81.0)	58.1 (39.1, 75.5)	67.7 (48.6, 83.3)	80.6 (61.9, 91.8)	71.0 (52.0, 85.8)	87.1 (70.2, 93.4)
	NPV	90.9 (75.8, 97.0)	85.7 (71.9, 93.4)	87.5 (75.2, 94.2)	92.6 (74.2, 98.7)	95.7 (78.4, 99.3)	100.0 (100.0, 100.0)
	PPV	26.7 (14.8, 43.2)	18.8 (8.6, 36.2)	23.1 (10.4, 43.7)	41.3 (13.3, 100.0)	35.7 (22.4, 51.7)	60.0 (37.5, 78.9)
	AUC	58.1 (26.2, 90.0)	66.9 (38.1, 95.7)	53.8 (24.5, 83.0)	84.1 (67.1, 100.0)	88.7 (68.2, 100.0)	97.3 (92.3, 100.0)

JRH

Abbreviations: NPV: Negative predictive value; PPV: Positive predictive value; AUC: The area under the receiver operator curve; CRP: C-reactive protein; LDH: Lactate dehydrogenase; NLR: Neutrophil-to-lymphocyte ratio; IL-6: Interleukin 6; CI: Confidence interval.

sons for decreasing lymphocytes in patients with severe COVID-19 [20]. Cytokines play critical roles in regulating immunological and inflammatory responses. IL-6 is a major inflammatory cytokine with pleiotropic effects implicated in coronavirus-induced storms [21, 22]. The cytokine storm, ie, high and uncontrolled levels of cytokines, triggers following the inflammatory responses induced in immune, epithelial, and endothelial cells due to SARS-CoV-2 invasion in the body [22]. Cardiovascular collapse and dysfunction syndromes of various organs, including renal and liver failure, are the consequences of high concentrations of the cytokine storm [23].

In this study, the bivariate analyses showed that disease progression was significantly associated with older age, consistent with other studies [24-26]. Older adults are more likely to develop more severe COVID-19 complications due to comorbidities, reduced physical functioning, poor body resistance, and a decline in angiotensin-converting enzyme-2 (ACE2) expression levels and anti-inflammatory response [4, 26].

Fever and respiratory distress were also associated with the risk factors of disease progression. Fever is the most frequent symptom in COVID-19 patients, which occurs in the immune system response to virus infections in the body [27]. Delay in seeking medical attention and the rapid disease progression in patients with COVID-19 could lead to an excessive inflammatory condition called a cytokine storm, which appears with an unlimited fever [28], promoting further inflammation and further immune activation with undesired effects [27]. Respiratory symptoms are also important indicators of the severity of the infection.

Similar to other studies [29, 30], patients with preexisting hypertension or chronic lung disease comorbidities were more susceptible to COVID-19 progression in the present study. Severe COVID-19 outcomes in patients with preexisting hypertension could be related to endothelial dysfunction and renin-angiotensin system (RAS) imbalance that favors a pro-inflammatory state, caus-

ing a higher level of IL-6 and tumor necrosis factor-α (TNF-α) [31, 32]. Elevating respiratory problems in COVID-19 patients with preexisting chronic lung diseases could worsen the condition and increase mortality [33].

Bivariate analyses demonstrated a significant association between elevated CRP and LDH values and disease progression in patients with COVID-19. In line with the findings of our study, numerous research [34-36] proposed that increased CRP and LDH values are related to higher levels of severity of COVID-19 and poorer outcomes. CRP is a plasma protein clinically used as a biomarker to identify disease severity in various inflammatory conditions [37]. COVID-19 progression may be accompanied by a cytokine storm, activation of the complement system, and amplifying inflammatory insults because of increasing CRP production through stimulating hepatocytes by cytokines such as IL-6 and TNFα. However, it is difficult to effectively state adaptive immunity in severe or critically ill COVID-19 patients due to severe damage to the integrity of the alveolar epithelial and endothelial barrier and significant decreases in lymphocyte counts with T cell-mediated immunosuppression. Therefore, it leads to severe macrophage infiltration and worsens acute lung injury [38]. LDH is an intracellular enzyme found in almost all body cells, and increasing its concentration may indicate damage to tissue/cells and viral infections or lung damage [39]. Nevertheless, the progress and prognosis of the pathogenic mechanism of LDH on COVID-19 remains unclear [40].

To date, corticosteroids and Remdesivir are the two most promising treatments for COVID-19 [41]. Nevertheless, as a result of the scarcity of randomized trials and inconclusive observational studies, the efficacy and safety of corticosteroids in viral pneumonia patients and Remdesivir's risk and benefit for COVID-19 patients who require high-flow oxygen or mechanical ventilation are not certain [42, 43]. In this study, taking Remdesivir in patients with poorer outcomes was statistically significantly less, and hydrocortisone was more.

Based on the multivariate logistic regression model results, IL-6, NLR, and lung involvement on admission were the only independent biomarkers associated with COVID-19 progression. Our risk prediction model with these three variables achieved an acceptable receiver operating characteristic (ROC)-AUC of 0.90 in the discovery dataset and 0.89 in the validation dataset for disease progression prediction. However, for the validation dataset, confidence intervals were wider for indices, especially for sensitivity, which may be related to the smaller sample size in the validation dataset. Therefore, a larger sample size is needed to assess its applicability in future research.

Conclusion

We identified a risk score that may represent a potential diagnostic tool in the clinical setting to identify COVID-19 patients at a higher risk of disease progression. This risk score model might have significant potential in public health. It can aid in making informed decisions and providing targeted interventions while maximizing efficiency and utilizing available resources. Considering other changes in biomarkers and clinical and diagnostic symptoms of patients with COVID-19 can help us determine the exact indicators of the disease progression.

Strengths and limitations

This study had some strengths. One of this study's strengths was the study's prospective nature, which added to the validity of data and findings. This study validated the risk score, and all indices for assessing the risk score performance were reported with a confidence interval. An additional strength was the simplicity of the achieved risk score. The study had some limitations, too. First, it was conducted in a single center. Second, in this study, the risk score model had an acceptable ROC-AUC in the discovery and validation datasets for disease progression prediction. However, for the validation dataset, confidence intervals were wider for indices, especially for sensitivity, which may be related to the smaller sample size in the validation dataset. Therefore, a single-center study and a small sample size weakened the generalizability of the study. Thus, future studies with national data and larger sample sizes are needed.

Ethical Considerations

Compliance with ethical guidelines

This study was approved by the Ethics Committee of Gonabad University of Medical Sciences (Code: IR.GMU.REC.1399.079). All phases of this study were based on the ethical principles of human medical research (Helsinki Declaration). Each patient or their respective families provided informed consent.

Funding

This study was conducted with support from the Social Development & Health Promotion Research Center, Gonabad University of Medical Sciences (No.: 139932).

Authors' contributions

Idea presentation: Mohsen Sahebanmaleki, Sina Sabeti Bilondi and Ali Delshad Noughabi; Study design and conceptualization: Fatemeh Mohammadzadeh, Mitra Tavakolizadeh, and Jafar Hajavi; Data collection: Mohsen Sahebanmaleki, Sina Sabeti Bilondi, Ali Delshad Noughabi and Hosein Aalami; Data analysis and writing the original draft: Fatemeh Mohammadzadeh; Review, editing and final Approval: All authors.

Conflict of interest

The authors declared no conflict of interest.

Acknowledgments

The authors gratefully acknowledge the Social Development and Health Promotion Research Center, Gonabad University of Medical Sciences, and all the people who participated in the study.

References

[1] Lu H, Stratton CW, Tang YW. Outbreak of pneumonia of unknown etiology in Wuhan, China: The mystery and the miracle. Journal of Medical Virology. 2020; 92(4):401-2. [DOI:10.1002/jmv.25678] [PMID]

[2] Cucinotta D, Vanelli M. WHO declares COVID-19 a pandemic. Acta Bio-Medica: Atenei Parmensis. 2020; 91(1):157-60. [PMID]

[3] Marcolino MS, Pires MC, Ramos LEF, Silva RT, Oliveira LM, Carvalho RLR, et al. ABC2-SPH risk score for in-hospital mortality in COVID-19 patients: development, external validation and comparison with other available scores. International Journal of Infectious Diseases. 2021; 110:281-308. [DOI:10.1016/j.ijid.2021.07.049] [PMID]

[4] Aboumrad M, Zwain G, Smith J, Neupane N, Powell E, Dempsey B, et al. Development and validation of a clinical risk score to predict hospitalization within 30 days of coronavirus disease 2019 diagnosis. Military Medicine. 2023; 188(3-4):e833-40. [DOI:10.1093/milmed/usab415] [PMID]

[5] WHO. Coronavirus (COVID-19) dashboard. Geneva: World Health Organization; 2023. [Link]

[6] Grant SW, Collins GS, Nashef SA. Statistical Primer: Developing and validating a risk prediction model. European Journal of Cardio-Thoracic Surgery. 2018; 54(2):203-8. [DOI:10.1093/ejcts/ezy180] [PMID]

[7] Janes H, Pepe MS, Gu W. Assessing the value of risk predictions by using risk stratification tables. Annals of Internal Medicine. 2008; 149(10):751-60. [DOI:10.7326/0003-4819-149-10-200811180-00009] [PMID]

[8] Levine DM, Lipsitz SR, Co Z, Song W, Dykes PC, Samal L. Derivation of a clinical risk score to predict 14-day occurrence of hypoxia, ICU admission, and death among patients with coronavirus disease 2019. Journal of General Internal Medicine. 2021; 36(3):730-7. [DOI:10.1007/s11606-020-06353-5] [PMID]

[9] Liang W, Liang H, Ou L, Chen B, Chen A, Li C, et al. Development and validation of a clinical risk score to predict the occurrence of critical illness in hospitalized patients with COVID-19. AMA Internal Medicine. 2020; 180(8):1081-9. [DOI:10.1001/jamainternmed.2020.2033] [PMID]

[10] Talebi A, Borumandnia N, Jafari R, Pourhoseingholi MA, Jafari NJ, Ashtari S, et al. Predicting the COVID-19 Patients Status Using Chest CT Scan Findings: A risk assessment model based on decision tree analysis. Advances in Experimental Medicine and Biology. 2023; 1412:237-50. [DOI:10.1007/978-3-031-28012-2_13] [PMID]

[11] Sharifi M, Khademian MH, Mousavi-Roknabadi RS, Ebrahimi V, Sadegh R. A new rapid approach for predicting death in coronavirus patients: The development and validation of the COVID-19 risk-score in fars province (CRSF). Iranian Journal of Public Health. 2022; 51(1):178-87. [DOI:10.18502/ijph.v51i1.8310] [PMID]

[12] Pazoki M, Payandemehr P, Montazeri M, Kafan S, Sheikhvatan M, Sotoodehnia M, et al. Risk stratification for in-hospital mortality in adult patients with COVID-19. Journal of Iranian Medical Council. 2022; 4(4):322-34. [DOI:10.18502/jimc.v4i4.8480]

[13] Amini N, Shalbaf A. Automatic classification of severity of COVID-19 patients using texture feature and random forest based on computed tomography images. International Journal of Imaging Systems and Technology. 2022; 32(1):102-10. [DOI:10.1002/ima.22679] [PMID]

[14] Nahm FS. Receiver operating characteristic curve: Overview and practical use for clinicians. Korean Journal of Anesthesiology. 2022; 75(1):25-36. [DOI:10.4097/kja.21209] [PMID]

[15] Jiang J. Applied medical statistics. New Jersey: John Wiley & Sons; 2022. [DOI:10.1002/9781119716822]

[16] Garcia-Gordillo JA, Camiro-Zúñiga A, Aguilar-Soto M, Cuenca D, Cadena-Fernández A, Khouri LS, et al. COVID-IRS: A novel predictive score for risk of invasive mechanical ventilation in patients with COVID-19. PLoS One. 2021; 16(4):e0248357. [DOI:10.1371/journal.pone.0248357] [PMID]

[17] Chandna A, Mahajan R, Gautam P, Mwandigha L, Gunasekaran K, Bhusan D, et al. Facilitating safe discharge through predicting disease progression in moderate coronavirus disease 2019 (covid-19): A prospective cohort study to develop and validate a clinical prediction model in resource-limited settings. Clinical Infectious Diseases. 2022; 75(1):e368-79. [DOI:10.1093/cid/ciac224] [PMID]

[18] Mantoo S, Shabir A, Khan UH, Shah TH, Siraj F, Mehfooz N, et al. Risk factors and a novel score (CARI 65) predicting mortality in COVID 19 patients. Indian Journal of Respiratory Care. 2022; 11:154-61. [DOI:10.4103/ijrc.ijrc_3_22]

[19] Halmaciu I, Arbăna-i EM, Kaller R, Mure-an AV, Arbăna-i EM, Bacalbasa N, et al. Chest CT severity score and systemic inflammatory biomarkers as predictors of the need for invasive mechanical ventilation and of covid-19 patients' mortality. Diagnostics. 2022; 12(9):2089. [DOI:10.3390/diagnostics12092089] [PMID]

[20] Rasyid H, Sangkereng A, Harjianti T, Soetjipto AS. Impact of age to ferritin and neutrophil-lymphocyte ratio as biomarkers for intensive care requirement and mortality risk in COVID-19 patients in Makassar, Indonesia. Physiological Reports. 2021; 9(10):e14876. [DOI:10.14814/phy2.14876] [PMID]

[21] Rostamian A, Ghazanfari T, Arabkheradmand J, Edalatifard M, Ghaffarpour S, Salehi MR, et al. Interleukin-6 as a potential predictor of COVID-19 disease severity in hospitalized patients and its association with clinical laboratory routine tests. Immunoregulation. 2020; 3(1):29-36. [DOI:10.32598/IMMUNOREGULATION.3.1.4]

[22] Han Q, Guo M, Zheng Y, Zhang Y, De Y, Xu C, et al. Current evidence of interleukin-6 signaling inhibitors in patients with COVID-19: A systematic review and meta-analysis. Frontiers in Pharmacology. 2020; 11:615972. [DOI:10.3389/fphar.2020.615972] [PMID]

[23] Niculet E, Chioncel V, Elisei AM, Miulescu M, Buzia OD, Nwabudike LC, et al. Multifactorial expression of IL6 with update on COVID19 and the therapeutic strategies of its blockade. Experimental and Therapeutic Medicine. 2021; 21(3):263. [DOI:10.3892/etm.2021.9693] [PMID]

[24] Booth A, Reed AB, Ponzo S, Yassaee A, Aral M, Plans D, et al. Population risk factors for severe disease and mortality in COVID-19: A global systematic review and meta-analysis. PloS One. 2021; 16(3):e0247461. [DOI:10.1371/journal.pone.0247461] [PMID]

[25] de Souza FSH, Hojo-Souza NS, Batista BDO, da Silva CM, Guidoni DL. On the analysis of mortality risk factors for hospitalized COVID-19 patients: A data-driven study using the major Brazilian database. PloS One. 2021; 16(3):e0248580. [DOI:10.1371/journal.pone.0248580] [PMID]

[26] Bajaj V, Gadi N, Spihlman AP, Wu SC, Choi CH, Moulton VR. Aging, immunity, and COVID-19: How age influences the host immune response to coronavirus infections? Frontiers in Physiology. 2021; 11:571416. [DOI:10.3389/fphys.2020.571416] [PMID]

[27] Gul MH, Htun ZM, Inayat A. Role of fever and ambient temperature in COVID-19. Expert Review of Respiratory Medicine. 2021; 15(2):171-3. [DOI:10.1080/17476348.2020.1816172] [PMID]

[28] Mehta P, McAuley DF, Brown M, Sanchez E, Tattersall RS, Manson JJ, et al. COVID-19: Consider cytokine storm syndromes and immunosuppression. Lancet. 2020; 395(10229):1033-4. [DOI:10.1016/S0140-6736(20)30628-0] [PMID]

[29] Wu Z, McGoogan JM. Characteristics of and important lessons from the coronavirus disease 2019 (COVID-19) outbreak in China: Summary of a report of 72 314 cases from the Chinese Center for Disease Control and Prevention. JAMA. 2020; 323(13):1239-42. [DOI:10.1001/jama.2020.2648] [PMID]

[30] Wendel Garcia PD, Fumeaux T, Guerci P, Heuberger DM, Montomoli J, Roche-Campo F, et al. Prognostic factors associated with mortality risk and disease progression in 639 critically ill patients with COVID-19 in Europe: Initial report of the international RISC-19-ICU prospective observational cohort. EClinicalMedicine. 2020; 25:100449. [DOI:10.1016/j.eclinm.2020.100449] [PMID]

[31] Muhamad SA, Ugusman A, Kumar J, Skiba D, Hamid AA, Aminuddin A. COVID-19 and hypertension: The what, the why, and the how. Frontiers in Physiology. 2021; 12:665064. [DOI:10.3389/fphys.2021.665064] [PMID]

[32] De Miguel C, Rudemiller NP, Abais JM, Mattson DL. Inflammation and hypertension: New understandings and potential therapeutic targets. Current Hypertension Reports. 2015; 17(1):507. [DOI:10.1007/s11906-014-0507-z] [PMID]

[33] Arman A, Tajik M, Nazemipour M, Ahmadinejad Z, Shahrestanaki SK, Hazrati E, et al. Risk factors of developing critical conditions in Iranian patients with COVID-19. Global Epidemiology. 2021; 3:100046. [DOI:10.1016/j.gloepi.2020.100046] [PMID]

[34] El-Shabrawy M, Alsadik ME, El-Shafei M, Abdelmoaty AA, Alazzouni AS, Esawy MM, et al. Interleukin-6 and C-reactive protein/albumin ratio as predictors of COVID-19 severity and mortality. The Egyptian Journal of Bronchology. 2021; 15(1):5. [DOI:10.1186/s43168-021-00054-1]

[35] Yildiz H, Castanares-Zapatero D, Pierman G, Pothen L, De Greef J, Aboubakar Nana F, et al. Validation of neutrophil-to-lymphocyte ratio cut-off value associated with high in-hospital mortality in covid-19 patients. International Journal of General Medicine. 2021; 14:5111-7. [DOI:10.2147/IJGM.S326666] [PMID]

[36] Broman N, Rantasärkkä K, Feuth T, Valtonen M, Waris M, Hohenthal U, et al. IL-6 and other biomarkers as predictors of severity in COVID-19. Annals of Medicine. 2021; 53(1):410-2. [DOI:10.1080/07853890.2020.1840621] [PMID]

[37] Kermali M, Khalsa RK, Pillai K, Ismail Z, Harky A. The role of biomarkers in diagnosis of COVID-19-A systematic review. Life Sciences. 2020; 254:117788. [DOI:10.1016/j.lfs.2020.117788] [PMID]

[38] Luan YY, Yin CH, Yao YM. Update advances on C-reactive protein in COVID-19 and other viral infections. Frontiers in Immunology. 2021; 12:720363. [DOI:10.3389/fimmu.2021.720363] [PMID]

[39] Bartziokas K, Kostikas K. Lactate dehydrogenase, COVID-19 and mortality. Medicina Clínica. 2021; 156(1):37. [DOI:10.1016/j.medcli.2020.07.043]

[40] Huang Y, Guo L, Chen J, Wu M, Zhang C, Liu Z, et al. Serum lactate dehydrogenase level as a prognostic factor for covid-19: A retrospective study based on a large sample size. Frontiers in Medicine. 2022; 8:671667. [DOI:10.3389/fmed.2021.671667] [PMID]

[41] Lee TC, McDonald EG, Butler-Laporte G, Harrison LB, Cheng MP, Brophy JM. Remdesivir and systemic corticosteroids for the treatment of COVID-19: A Bayesian re-analysis. International Journal of Infectious Diseases. 2021; 104:671-6. [DOI:10.1016/j.ijid.2021.01.065] [PMID]

[42] Dequin PF, Heming N, Meziani F, Plantefève G, Voiriot G, Badié J, et al. Effect of hydrocortisone on 21-day mortality or respiratory support among critically Ill patients with COVID-19: A randomized clinical trial. JAMA. 2020; 324(13):1298-306. [DOI:10.1001/jama.2020.16761] [PMID]

[43] Young B, Tan TT, Leo YS. The place for remdesivir in COVID-19 treatment. The Lancet. Infectious Diseases. 2021; 21(1):20-1. [DOI:10.1016/S1473-3099(20)30911-7] [PMID]

This Page Intentionally Left Blank