

Access this article online
Quick Response Code:

Website: www.jehp.net
DOI: 10.4103/jehp.jehp_1017_21

Using logistic regression to develop a diagnostic model for COVID-19: A single-center study

Raof Nopour¹, Mostafa Shanbehzadeh², Hadi Kazemi-Arpanahi^{3,4}

Abstract:

BACKGROUND: The main manifestations of coronavirus disease-2019 (COVID-19) are similar to the many other respiratory diseases. In addition, the existence of numerous uncertainties in the prognosis of this condition has multiplied the need to establish a valid and accurate prediction model. This study aimed to develop a diagnostic model based on logistic regression to enhance the diagnostic accuracy of COVID-19.

MATERIALS AND METHODS: A standardized diagnostic model was developed on data of 400 patients who were referred to Ayatollah Tallegghani Hospital, Abadan, Iran, for the COVID-19 diagnosis. We used the Chi-square correlation coefficient for feature selection, and logistic regression in SPSS V25 software to model the relationship between each of the clinical features. Potentially diagnostic determinants extracted from the patient's history, physical examination, and laboratory and imaging testing were entered in a logistic regression analysis. The discriminative ability of the model was expressed as sensitivity, specificity, accuracy, and area under the curve, respectively.

RESULTS: After determining the correlation of each diagnostic regressor with COVID-19 using the Chi-square method, the 15 important regressors were obtained at the level of $P < 0.05$. The experimental results demonstrated that the binary logistic regression model yielded specificity, sensitivity, and accuracy of 97.3%, 98.8%, and 98.2%, respectively.

CONCLUSION: The destructive effects of the COVID-19 outbreak and the shortage of healthcare resources in fighting against this pandemic require increasing attention to using the Clinical Decision Support Systems equipped with supervised learning classification algorithms such as logistic regression.

Keywords:

Coronavirus, coronavirus disease-2019, logistic regression, prognostic modeling

Introduction

Since December 2019, a new strand of coronavirus disease at first named severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) or 2019 n-CoV has emerged in Wuhan District, Hubei Province, China. It is thought that SARS-CoV-2 has animal origins that slipped from animal species into the human population.^[1,2] The World Health Organization later on February 11, 2020 announced coronavirus disease

2019 "COVID-19" as the name of this new disease.^[3,4] COVID-19 is a highly contagious viral infectious disease and continues to spreading aggressively around the whole world. The virus affects the lungs and causes severe respiratory pneumonia and is more dangerous in people with underlying conditions such as low immune function system, cardiovascular and respiratory diseases, malignancies, co-infectious diseases, and diabetes.^[5-10] The complex and highly contagious nature of the COVID-19 had led the WHO to declare this outbreak a

This is an open access journal, and articles are distributed under the terms of the Creative Commons Attribution-NonCommercial-ShareAlike 4.0 License, which allows others to remix, tweak, and build upon the work non-commercially, as long as appropriate credit is given and the new creations are licensed under the identical terms.

For reprints contact: WKHLRPMedknow_reprints@wolterskluwer.com

How to cite this article: Nopour R, Shanbehzadeh M, Kazemi-Arpanahi H. Using logistic regression to develop a diagnostic model for COVID-19: A single-center study. *J Edu Health Promot* 2022;11:153.

¹Department of Health Information Management, Student Research Committee, School of Health Management and Information Sciences Branch, Iran University of Medical Sciences, Tehran, Iran, ²Department of Health Information Technology, School of Paramedical, Ilam University of Medical Sciences, Ilam, Iran, ³Department of Health Information Technology, Abadan University of Medical Sciences, Abadan, Iran, ⁴Student Research Committee, Abadan University of Medical Sciences, Abadan, Iran

Address for correspondence:

Dr. Hadi Kazemi-Arpanahi, Department of Health Information Technology, Abadan University of Medical Sciences, Abadan, Iran.
E-mail: h.kazemi@abadanums.ac.ir

Received: 10-07-2021
Accepted: 25-08-2021
Published: 11-06-2022

public health emergency, which consequently brought significant health, economic, and social challenges.^[11,12] It is mandatory to be able to differentiate between COVID-19 from other pneumonia-like diseases early after symptom development.^[4] Due to the high-level spread and increasing epidemiology trend of COVID-19, its early diagnosis and rapid isolation of infected people play a key role in confining this virus and thereby reducing the disease outbreak and mortality rate.^[9] Deadly complications, long latency period, difficulties of detection and testing, unknown of many characteristics, ambiguous transmission modes, and differential diagnosis with other respiratory morbidities such as pneumonia, have increased the vagueness and challenges in understanding the risk factors leading to the disease, its natural history, and controlling the outbreak.^[13] In addition, it is necessary to seek early detection and isolation of positive cases as rapidly and accurately as possible for containing the transmission of the virus, especially for asymptomatic cases in an early stage.^[4]

Now, accurate nucleic acid detection plays a key role in COVID-19 detection and inhibition, and reverse transcription-polymerase chain reaction is the core technique for nucleic acid detection. However, it is not possible to use this time-consuming and labor-intensive method for screening a large increasing number of supposed individuals and symptomless infected cases.^[14] Thus, an accurate and reliable diagnostic method to provide patient risk classification to support reliable clinical decision-making with the hope of improving patient outcomes and quality of care can help reduce misdiagnosis and poor prognosis of COVID-19.^[13] Therefore, to fight against COVID-19, risk assessment has become increasingly imperative during this pandemic. Several different statistical methods can be used to develop a risk prediction model including but not restricted to logistic regression, linear regression, Cox regression, and machine learning (ML).^[15-17] To address the uncertainties in COVID-19 diagnosis, logistic regression as a supervised learning classification algorithm can be utilized to provide patient risk stratification to support tailored clinical decision-making including measuring the probability of a disease, assessing the disease likelihood, forecasting the spread, and predicting fatality.^[13,18-21] To study the risk factors associated with COVID-19, logistic regressions have become a fundamental section of any data analysis related to the explanation of association among an outcome variable and one or more predictor variables. Therefore, this study was undertaken to develop a diagnostic model to predict the risk of the development of COVID-19. This can be used as a quick screening tool to improve the diagnostic efficiency of COVID-19 through statistical analysis of the significant factors.

Materials and Methods

Study design and setting

A hospital-based, retrospective, and applied study was designed in Ayatollah Taleghani hospital (COVID-19 referral center) Abadan, Iran, to develop a diagnosis model using logistic regression analysis in COVID-19 patients.

Study participants and sampling

The study participants included case records of 435 patients that referred to the center for COVID-19 diagnosis. Of the 435 case records for the study, 35 cases of missing data, have been excluded from the analysis. Among participants who were eligible for the study, 250 cases belonged to people with laboratory-confirmed infection, and 150 cases belonged to NonCOVID-19 patients suffering from pneumonia of other origins.

Model development and evaluation

For model development, multiple analyses were conducted to extract the most important regressors from the patient's medical record where COVID-19 positive cases were classified as 1 and healthy individuals classified as 0. For preliminary analysis of the data, the Chi-Square correlation coefficient method was used. Then, the diagnostic regressors that are correlated with the output class variable (in COVID-19-positive and non-COVID-19 groups) was obtained at the level of $P < 0.05$ were selected. Furthermore, to create a suitable model for COVID-19 diagnosis, by considering the output class (binary), the binary logistic regression (BLR) with the backward conditional method was used. In the current logistic regression model, the success was defined when the dependent variable took the value positive COVID-19. Finally, the efficiency of the developed model was evaluated using the amount of Log-likelihood (Formula1), confusion matrix, as well as its accuracy, precision, and sensitivity were determined.

The amount of Log-likelihood is a coefficient of the observed and predicted value according to Equation (1) and reducing it by adding the variables indicates a better model performance. In Table 1, the true positive (TP) and true negative (TN) represent the number of samples that belonged to sick and healthy individuals, which are correctly classified by the model, respectively. The false positive (FP) represents the number of healthy individuals that are classified as the patient, falsely by the model. The false negative (FN) represents the patients that the model is incorrectly classified as healthy people, therefore, the specificity, sensitivity, and accuracy of the model, according to relationships were obtained (Equation 1-3).

Table 1: If term removed model (step 5)

Variable name	Model log-likelihood	Change in -2 log-likelihood	Df	Significance of change
History of contact with suspected people	-21.441	5.527	1	0.19
History of alcohol consumption	-19.022	0.668	1	0.407
History of the immunosuppressive drug in taking	-18.849	0.342	1	0.552
History of ARDS	-19.280	1.203	1	0.273
Fever	-28.478	19.601	1	0
Cough	-20.915	4.473	1	0.034
Chest pain	-19.452	1.548	1	0.213
Disability	-19.649	1.941	1	0.164
Rhinorrhoea	-22.991	8.626	1	0.003
Lung lesion	-51.455	65.554	1	0

ARDS=Acute respiratory distress syndrome, Df=Degree of freedom

$$\text{Specificity} = \frac{TN}{TN + FP} \quad (1)$$

$$\text{Sensitivity} = \frac{TP}{TP + FN} \quad (2)$$

$$\text{Accuracy} = \frac{TN + TP}{TN + FP + TN + TP} \quad (3)$$

The area under the curve

The area under the curve (also called AUC) is equal to the probability that one algorithm will score a selected positive sample higher than a selected negative one in a random state. The area under the curve is given by (the integral boundaries are reversed as large T has a lower value on the x-axis) (Equations 4 and 5).^[22]

$$A = \int_{z=0}^1 TPR(FPR^{-1}(\chi))d\chi = \int_{-\infty}^{\infty} TPR(T)FPR'(T)dT = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} I(T' > T)f_1(T')f_0(T)dT'dT = p(X_1 > X_0) \quad (4)$$

Equation 5:

$$\text{log-likelihood} = \sum_{i=1}^N (Y_i | n[P(Y_i)] + [1 - Y_i] | n[1 - P(Y_i)]) \quad (5)$$

Ethical consideration

The research deputy of Abadan University of Medical Sciences (ethical code: IR.ABADANUMS.REC.1400.064) approved the current study. To protect the privacy and confidentiality of patients, we concealed the unique identifying information of all patients in the process of data collection and presentation.

Results

After analyzing all the diagnostic regressors of COVID-19 (or predictors) using the Chi-square bivariate correlation method, the 15 important variables were identified, as follows:

Blood oxygen saturation, with $P < 0.01$ was obtained

as a useful laboratory finding in the diagnosis of COVID-19. The history of alcohol consumption ($P = 0.01$), consumption of immunosuppressive drugs ($P < 0.01$), history of respiratory failure (acute respiratory distress syndrome) ($P < 0.01$), history of respiratory infection (pneumonia) ($P = 0.01$) were considered as important diagnostics factors.

The history of traveling to high-risk areas ($P < 0.01$) and the history of contact with people with the suspected COVID-19 ($P < 0.01$) were considered effective epidemiological factors in diagnosing COVID-19.

The developing COVID-19 diagnostic model based on binary logistic regression analysis using the conditional backward method was performed in five steps that are shown in Table 1.

In this method, by removing the diagnostic regressors of COVID-19 disease in five steps (each step equals the removal of one variable) that were less important [Table 1], the 10 diagnostic regressors that are the most important variables in creating a regression model for diagnosing COVID-19 disease has been determined.

In this model, by removing the less important variables, the log-likelihood function was reduced for a certain degree of freedom (DF = 1); therefore, the efficiency of the logistic regression model in diagnosing COVID-19 has been improved by five sequential steps.

Figures 1 and 2 show the results of the classification of possible groups predicted by the regression detection model in terms of different frequency values observed at the 1st and 5th steps of this model, respectively.

The criterion for classifying healthy and sick people according to the predicted value (cutoff value) in these forms, as can be seen, is 0.5, in other words, the value is higher than 0.5 in the group of sick people and lower than 0.5 in the group of healthy people. Furthermore, each number (case) in the figure represents five samples. As shown in Figure 1, a case representing COVID-19 in the

iterated frequency of 10 was incorrectly classified in the predicted probability group ($0 < X < 0.1$) (X = predicted probability group). However, at the 5th step of the regression model, all cases were classified correctly.

Furthermore, in the 5th step, all the cases predicted by the model were in according to the actual cases for different iterated frequencies, which has a better performance than the first step. Furthermore, the cases without disease are better classified in the 5th step than the 1st step with an incorrectly classified case described. The result of classifying the cases of the sick and healthy individuals at the 1st and 5th steps of classifying of BLR function is shown in Table 2.

Based on the comparison of the information obtained from Figures 1 and 2, it is observed that by removing the less important variables than the regression model at the 5th step compared to the first step, the amount of TP and FN (number of correctly classified items) obtained from the last step, as the Figure 2 shows have been increased compared to the first. In the 1st step, the percentage of the correctly classified cases was 78% and in the 5th, the percentage of the correctly classified cases was 98%.

Therefore, by omitting diagnostic criteria of less importance than the model, a significant improvement in its efficiency has been achieved and the percentage of

correctly classified cases has increased by 20% as shown in Figure 3. The confusion matrix in Table 2 (step 5), and Table 2 shows the specificity, sensitivity, and accuracy of 97.3%, 98.8%, and 98.2%, respectively.

The result of the receiver operator characteristics (ROC) at all five steps is shown in Figure 4. The vertical and horizontal vertices show TP rate (TPR) and FP rate, respectively.

Based on Figure 4, in the 5th step, the ROC curve was closer to the TPR than in the initial steps, so in this step, it had better performance in classifying the true cases and generally had better performance in the classification of study samples.

Table 2: The 1th and 5th step of binary logistic regression confusion matrix

BLR's steps	Cases classified
1 st step	
197	35
53	115
5 th step	
247	4
3	146

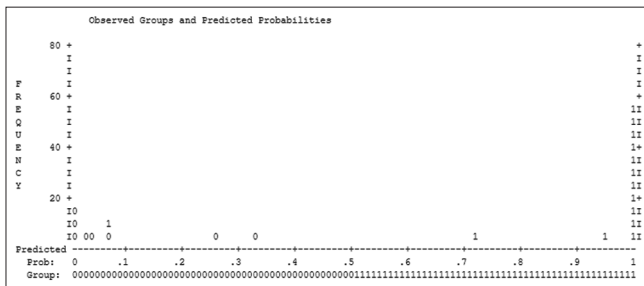


Figure 1: The predicted probability group according to the different iteration frequency (step 1)

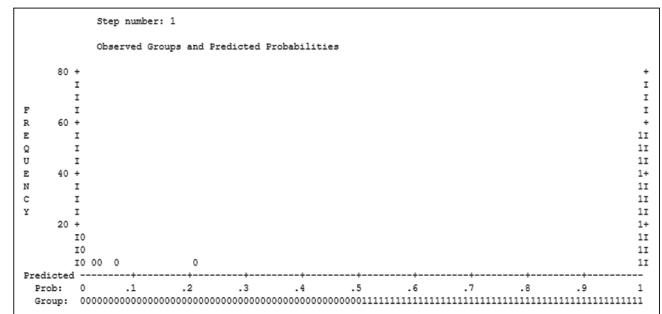


Figure 2: The predicted probability group according to the different iteration frequency (step 5)

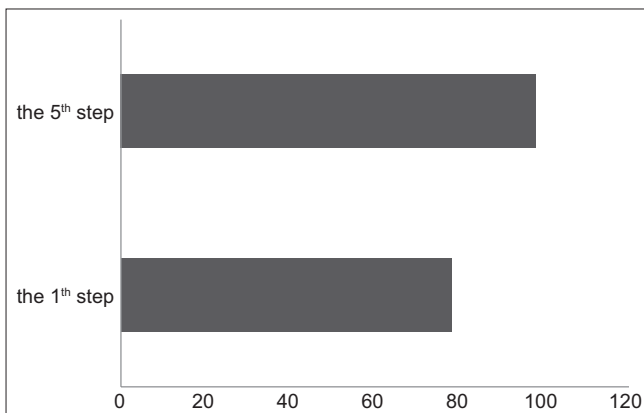


Figure 3: Comparison of the BLR values

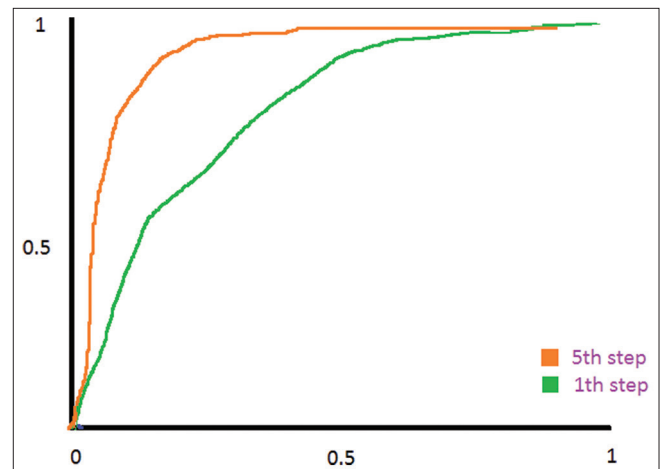


Figure 4: The receiver operator characteristics graph of BLR in five steps

Discussion

The exponential outbreak of the COVID-19 and the continuously increased number of infected cases without effective treatment forces the medical fraternity to discover the many unfamiliar aspects of the disease.^[23,24] In this situation, the health-care industry is imposed with great pressure with severe shortages of intensive care resources.^[25,26] Therefore, developing an accurate diagnostic model that can effectively predict COVID-19 presence (diagnosis) with important prognostic determinants is indeed vital.^[1,2] This study proposed an understandable, intuitive, and yet accurate prediction model using logistic regression based on the most important predictors.

The logistic regression model is a chief technique to recognize the principle that the goal of an analysis is the same as that of the traditional model building technique used in statistical theory to find a suitable description of the relationship between the outcome variable and predictor variables. Logistic regression was employed to assess hypotheses about the associations of the response variable with explanatory variables. It does not need normally distributed data compared with discriminant analysis.^[4,19] Logistic regression aids one to forecast the discrete outcome from a variety of variables. In the logistic regression model, we consider the outcome variable is a categorical random variable, getting only two likely outcomes named binary or dichotomous. The logistic regression allows physicians and researchers to recognize significant contributing factors related to COVID-19. This method also permits researchers to evaluate the extent of the effect of factors.^[27,28]

Some efforts have been focused on predicting COVID-19 using a regression model and regularly collected clinical data. Almeshal *et al.* devised compartmental and LR models to predict the spread of COVID-19 in Kuwait, their study revealed the high performance of the LR model to predict the peak of daily cases, the total infected cases, and the expected dates of the start and ending phase of the epidemic.^[13] Zhou *et al.* used ML based on logistic regression to predict the fatality rate of COVID-19 patients that can effectively predict the outcomes of COVID-19 patients with fatality probabilities (accumulative f1-score = 93.76% and accuracy score = 93.92%).^[29] Xu *et al.* applied ordinal logistic regression analysis to identify the determinants of illness severity of COVID-19.^[21] Bhandari *et al.* developed a predictor model of mortality risk in COVID-19 patients from routine hematologic parameters. The performance metrics of the model with 5-fold cross-validation showed an area under the receiver operating characteristic curve, sensitivity, specificity, and validation accuracy to be 0.95, 90%, 92%, and 70%, respectively.^[19] Medina-Mendieta *et al.* used Logistic Regression and Gompertz Curves to

predict peaks and total numbers of infected cases and deaths due to COVID-19. Both models showed good fit, low mean square errors, and all parameters were highly significant.^[20] Hu and Li showed that the AUC, sensitivity, and specificity of a death prediction model based on a logistic regression model for predicting the mortality of COVID-19 cases during hospitalization were 0.804, 83.8%, and 82.3%.^[30]

The novelty of this study lies in the use of logistic regression in developing a diagnostic model, it is extremely vital to develop an accurate diagnostic model that can effectively diagnose COVID-19 patients with important prognostic criteria. Designing a scientific and valid LR-based prediction model, while not a substitute for clinical experience, can assist in the early detection and effective supportive intervention to improve patient outcomes and quality of care and ultimately a decrease of death in COVID-19 patients.^[28,31] This led to decreasing ambiguity by offering quantitative, objective, and evidence-based models for risk stratification, prediction, and eventually episode of care planning.^[32]

Limitations and recommendations

This study had some limitations. First, we analyze a retrospective dataset that may lack control over data fields or incomplete data. Second, the dataset was extracted from a single hospital with a low sample size based on 400 data in one city, consequently, they may not be generalizable. In the future, the inclusion of other clinical and radiological features could contribute to increasing the accuracy of the prediction model.

Conclusion

In this study, we demonstrated the use of the logistic regression model to diagnose COVID-19 ahead of time. This method has many potential applications to provide frontline clinicians with an objective instrument to manage COVID-19 patients more efficiently in such time-sensitive, resource-demanding, and potentially resource-constrained situations. Finally, we believed further investigations are needed to validate our model in a larger, multi-central, and more qualitative dataset.

Acknowledgment

We thank the research deputy of the Abadan University of Medical Sciences for financially supporting this project IR.ABADANUMS.REC.1400.064.

Financial support and sponsorship

Nil.

Conflicts of interest

There are no conflicts of interest.

References

1. Li J, Li S, Qiu X, Zhu W, Li L, Qin B. Performance of diagnostic model for differentiating between COVID-19 and influenza: A 2-center retrospective study. *Med Sci Monit* 2021;27:e932361.
2. Varotsos CA, Krapivin VF, Xue Y. Diagnostic model for the society safety under COVID-19 pandemic conditions. *Saf Sci* 2021;136:105164.
3. Kumar R, Singh V, Mohanty A, Bahurupi Y, Gupta PK. Corona health-care warriors in India: Knowledge, attitude, and practices during COVID-19 outbreak. *J Educ Health Promot* 2021;10:44.
4. Shanbehzadeh M, Kazemi-Arpanahi H, Mazhab-Jafari K, Haghiri H. Coronavirus disease 2019 (COVID-19) surveillance system: Development of COVID-19 minimum data set and interoperable reporting framework. *J Educ Health Promot* 2020;9:203.
5. Liu Y, Wang Z, Ren J, Tian Y, Zhou M, Zhou T, *et al.* A COVID-19 risk assessment decision support system for general practitioners: Design and development study. *J Med Internet Res* 2020;22:e19786.
6. Alom MZ, Rahman MM, Nasrin MS, Taha TM, Asari VK. COVID_MNet: COVID-19 detection with multi-task deep learning approaches. *arXiv preprint* 2020:1-11. [arXiv:2004.03747](https://arxiv.org/abs/2004.03747).
7. Bansal M. Cardiovascular disease and COVID-19. *Diabetes Metab Syndr* 2020;14:247-50.
8. Lai CC, Shih TP, Ko WC, Tang HJ, Hsueh PR. Severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) and coronavirus disease-2019 (COVID-19): The epidemic and the challenges. *Int J Antimicrob Agents* 2020;55:105924.
9. Hussain A, Bhowmik B, do Vale Moreira NC. COVID-19 and diabetes: Knowledge in progress. *Diabetes Res Clin Pract* 2020;162:108142.
10. Moujaess E, Kourie HR, Ghosn M. Cancer patients and research during COVID-19 pandemic: A systematic review of current evidence. *Crit Rev Oncol Hematol* 2020;150:102972.
11. Cascella M, Rajnik M, Aleem A, Dulebohn SC, Di Napoli R. Features, evaluation, and treatment of coronavirus (COVID-19). *Statpearls* [internet]. 2022.
12. Sohrabi C, Alsafi Z, O'Neill N, Khan M, Kerwan A, Al-Jabir A, *et al.* World Health Organization declares global emergency: A review of the 2019 novel coronavirus (COVID-19). *Int J Surg* 2020;76:71-76.
13. Almeshal AM, Almazrouee AI, Alenizi MR, Alhajeri SN. Forecasting the spread of COVID-19 in Kuwait using compartmental and logistic regression models. *Appl Sci* 2020;10:3402.
14. Mardani R, Vasmehjani AA, Zali F, Gholami A, Nasab SD, Kaghazian H, *et al.* Laboratory parameters in detection of COVID-19 patients with positive RT-PCR; a diagnostic accuracy study. *Arch Acad Emerg Med* 2020;8:e43.
15. Gao Y, Cai GY, Fang W, Li HY, Wang SY, Chen L, *et al.* Machine learning based early warning system enables accurate mortality risk prediction for COVID-19. *Nat Commun* 2020;11:5033.
16. Kanagarathinam K, Sekar K. Estimation of the reproduction number and early prediction of the COVID-19 outbreak in India using a statistical computing approach. *Epidemiol Health* 2020;42:e2020028.
17. Amirhajlou L, Sohrabi Z, Alebouyeh MR, Tavakoli N, Haghghi RZ, Hashemi A, *et al.* Application of data mining techniques for predicting residents' performance on pre-board examinations: A case study. *J Educ Health Promot* 2019;8:108.
18. Singh H, Dhar J. Mathematical population dynamics and epidemiology in temporal and spatio-temporal domains. CRC Press; 2018.
19. Bhandari S, Shaktawat AS, Tak A, Patel B, Shukla J, Singhal S, *et al.* Logistic regression analysis to predict mortality risk in COVID-19 patients from routine hematologic parameters. *Ibnosina J Med Biomed Sci* 2020;12:123.
20. Medina-Mendieta JF, Cortés-Cortés M, Cortés-Iglesias M. COVID-19 forecasts for cuba using logistic regression and gompertz curves. *MEDICC Rev* 2020;22:32-9.
21. Xu K, Zhou M, Yang D, Ling Y, Liu K, Bai T, *et al.* Application of ordinal logistic regression analysis to identify the determinants of illness severity of COVID-19 in China. *Epidemiol Infect* 2020;148:e146.
22. Fawcett T. An introduction to ROC analysis. *Pattern Recognit Lett* 2006;27:861-74.
23. Koffman J, Gross J, Etkind SN, Selman L. Uncertainty and COVID-19: How are we to respond? *J R Soc Med* 2020;113:211-6.
24. Koffman J, Gross J, Etkind SN, Selman LE. Clinical uncertainty and COVID-19: Embrace the questions and find solutions. *Palliat Med* 2020;34:829-31.
25. Dijkman WM, van Acht NM, van Akkeren JP, Bhagwanbali RC, van Pul C. Comparing Ventilation parameters for COVID-19 patients using both long-term ICU and anesthetic ventilators in times of shortage. *J Intensive Care Med* 2021;36:963-71.
26. Kanji S, Burry L, Williamson D, Pittman M, Dubinsky S, Patel D, *et al.* Therapeutic alternatives and strategies for drug conservation in the intensive care unit during times of drug shortage: A report of the Ontario COVID-19 ICU drug task force. *Can J Anaesth* 2020;67:1405-16.
27. Javali SB, Sunkad MA, Wantamutte AS. Prediction of risk factors of periodontal disease by logistic regression: A study done in Karnataka, India. *Int J Community Med Public Health* 2018;5:5301-9.
28. Hu H, Du H, Li J, Wang Y, Wu X, Wang C, *et al.* Early prediction and identification for severe patients during the pandemic of COVID-19: A severe COVID-19 risk model constructed by multivariate logistic regression analysis. *J Glob Health* 2020;10:020510.
29. Zhou Y, He Y, Yang H, Yu H, Wang T, Chen Z, *et al.* Exploiting an early warning Nomogram for predicting the risk of ICU admission in patients with COVID-19: A multi-center study in China. *Scand J Trauma Resusc Emerg Med* 2020;28:106.
30. Hu K, Li B. Logistic regression analysis of death risk factors of patients with severe and critical coronavirus disease 2019 and their predictive value. *Zhonghua Wei Zhong Bing Ji Jiu Yi Xue* 2020;32:544-7.
31. Josephus BO, Nawir AH, Wijaya E, Moniaga JV, Ohyver M. Predict mortality in patients infected with COVID-19 virus based on observed characteristics of the patient using logistic regression. *Procedia Comput Sci* 2021;179:871-7.
32. Mohammadi F, Pourzamani H, Karimi H, Mohammadi M, Mohammadi M, Ardalan N, *et al.* Artificial neural network and logistic regression modelling to characterize COVID-19 infected patients in local areas of Iran. *Biomed J* 2021;44:304-16.