Explanation of COVID-19 mortality using artificial neural networks (ANN) based on underlying and laboratory risk factors in Ilam, Iran

Article · October 2021

DOI: 10.22092/ARI.2021.355438.1684

Gholamreza Kalvandi
Ilam University of Medical Sciences

Elham Shafiei
Ilam University of Medical Sciences

Archives of Razi Institute

DOI: 10.22092/ARI.2021.355438.1684

Explanation of COVID-19 Mortality Using Artificial Neural Network (ANN) Based on Underlying and Laboratory Risk Factors in Ilam, Iran

Taghinezhad, F[1], Kaffashian, M[2], Kalvandi, Gh[3], Shafiei, E[4]*

1. Clinical Research Development Unit, Mostafa Khomeini hospital, Ilam University of Medical Sciences, Ilam, Iran
2. Department of Physiology, School of Medicine, Ilam University of Medical Sciences, Ilam, Iran
3. Department of Pediatrics, School of Medicine, Besat Hospital, Hamadan University of Medical Sciences, Hamadan, Iran
4. Non-Communicable Center, Ilam University of Medical Sciences, Ilam, Iran

*Corresponding Author: Shafiei-e@medilam.ac.ir

Abstract

The spread of new waves of coronavirus outbreaks, high mortality rates, and time-consuming and numerous challenges in achieving collective safety through vaccination and the need to prioritize the allocation of vaccines in the general population have led to continued identification of risk factors associated with mortality in patients through innovative strategies and new statistical models. In this study, an artificial neural network (ANN) model was used to predict morbidity in patients with COVID-19. Data of 2206 patients were extracted from the registry program of Shahid Mostafa Khomeini Hospital in Ilam, Iran, and were randomly analyzed in two training (1544) and testing (662) groups. By fitting different models of a three-layer neural network, 12 variables can explain more than 77% of the mortality variance in COVID-19 patients. These findings could be used to better manage mortality, prioritize vaccinations, public education, and quarantine, and allocate intensive care beds to reduce COVID-19 mortality. The findings also confirmed the power of a better explanation of ANN models to predict the mortality of patients.

Keywords: artificial neural networks COVID-19 (Coronavirus); multilayer perceptron; Iran

1. Introduction

COIVD-19, as the greatest crisis of the health system and the challenge of human society since World War II, has severely affected health systems, the global economy, and governments, causing widespread and unimaginable deaths (1). Moreover, unprecedented investments were made in the prevention, diagnosis, and treatment of the disease, reducing the incidence rate and finally the mortality rate (2).

Undoubtedly, the greatest success against the recent pandemic can be attributed to the development of vaccines (3), a factor that raised great hopes for the eradication of the disease. Currently, several different vaccination projects have been able to obtain emergency use licenses after passing the phases to be used for general vaccination (4). However, numerous challenges and obstacles, including the level and amount of production, reasonable prices for the purchase and maintenance particularly in low- and middle-income countries, proper and fair allocation, doubts, and incomplete acceptance of vaccination have currently hampered the herd safety and control of the disease (5, 6). On the other hand, the arrival of new waves of disease outbreaks and widespread mortality, along with the lack of adequate vaccines in countries such as Iran have exacerbated the increasingly urgent situation. This has prioritized attempts to prevent the spread of the disease through reducing the incidence and mortality by identifying related risk factors (7, 8). On the other hand, the limited number of available vaccines compared to the high number of applicants has led to the identification of high-risk individuals in terms of risk of infection and mortality using low-cost contextual and laboratory variables to prioritize vaccination to an important topic in the health policy-making (9).

At present, there is no definitive and golden method that can predict the mortality rate of COVID-19 patients in terms of laboratory variables and underlying diseases. Creative and scientific study ideas for better management are a challenge to control the incidence and mortality rates (2). One of the creative strategies is to use accurate statistical and mathematical methods(10) in explaining and predicting the importance of each risk factor and experimental changes in the incidence, prevalence, and mortality of patients (11). Predicting the future situation based on existing variables will lead to better management and reduced costs when facing the crisis (12). In the field of COVID-19 crisis management, understanding the relationship between different factors and variables, and predicting events, such as mortality, are of interest to many researchers and health policy-makers and can pave the way for better decision-making in this area. Since the outbreak of the disease, clinical and laboratory factors, including changes in patients' lung CT scan, changes in hematology tests (lymphopenia, Thrombocytopenia, leukocytosis), changes in some biochemical tests (changes in liver enzymes and increases in troponin, urea, and creatinine), coagulation indicators (increase in PT time, D_dimer changes), and an increase in inflammatory indicators (increased ESR, CRP-

10), have been extensively studied in various studies in addition to the COVID-19 diagnostic test by PCR (as the gold standard of diagnosis) (13, 14).

The artificial neural network (ANN) model is an example of such non-classical models that can be used without any presuppositions, which is a kind of progress and transformation in various sciences, including medical sciences (15). The successful use of ANNs in many applications has been demonstrated in various articles in recent years. Practical examples include the use of this network in studies related to breast cancer and pancreatitis patients, explaining and predicting the number of hospitalization days in heart patients, explaining the possible effects of drugs on cancer patients, determining factors affecting hospital readmission, explaining clinical results after prostate cancer chemotherapy, etc. (16, 17). ANNs are new information processing structures that use methods specific to biological neural networks. The main difference between the ANN and other common methods is that there is no need in this method to consider a predetermined relationship between primary and secondary variables (18). Instead, these networks themselves discover the relationship between variables during the learning process. Artificial intelligence networks show their strength, especially in cases where there is a complex and nonlinear relationship between primary and secondary variables (19). As mentioned earlier in the use of classical models, they also require the existence of special assumptions, and the appropriateness of the model depends on the accuracy of these assumptions. Therefore, in cases where these assumptions do not seem correct using real data, it seems necessary to develop models that do not require such assumptions (20). ANN models have comparative advantages over other conventional methods. The need for theory and hypothesis, hypersensitivity to noise and error, the need for data completeness, inability to understand, track, and record relationships between multiple data are some of the limitations and problems with conventional and classical methods, which can be modified to a large extent in ANN models (21).

According to studies, many factors can contribute to the poor prognosis and mortality of COVid-19 patients. On the other hand, most of the information about the epidemiological and laboratory factors of COVID-19 is still incomplete and doubtful and there is insufficient evidence and certainty for their use. Recognizing the importance of each of the indicators related to deaths from disease using efficient output models can be considered in the process of better treatment and care, management of scarce resources during the peak of the disease, prioritization of special and hospitalization beds, better human resource management in the health sector, increasing the reliability of patients with lower risk for home care, and appropriate planning to prioritize the allocation of vaccines to higher risk groups. Therefore, this study aimed to explain the role of underlying and laboratory risk factors in the mortality of COVID-19 patients using an ANN model.

2. Material and Methods

2.1. Sample Collection

In the present retrospective cohort study, 21 field and laboratory variables were examined in 2206 patients with confirmed PCR and definitive COVID-19 diagnosis who were referred to Mostafa Khomeini Hospital in Ilam with a definitive diagnosis of COVID-19 from March 1, 2020 to March 20, 2021, and included in the study.

2.2. Input and out pout layer in ANN& logistic regression

Then, the strength of ANNs and logistic regression models in explaining the mortality caused by COVID-19 was compared with the ability to explain logistic regression based on contextual and laboratory variables. Participants were randomly divided into two training and testing groups. Input variables, namely diabetes, BMI, age, WBC, PLT, ESR, intubation, HTN, COPD, heart disease, hemoglobin, and cancer, were selected based on previous studies and preliminary analysis on the data of this study. The output variable was the patient's (deceased and non-deceased) condition. To evaluate the predictive power, different ANN models were compared with the logistic regression. The variables were arranged in descending order in terms of their effects on the output variables in a stepwise logistic regression model. A box plot was used to identify these cases in nominal variables, and leverage methods were used in numerical variables. The loss weighting method was used to obtain a robust and outlier-resistant regression model.

2.3.ANN analysis

Two training and testing groups were used to produce ANNs. First, multilayer perceptrons (MLPs) with a hidden layer and a neuron in the outer layer were trained in the first group. After obtaining the most effective possible combination to predict the output variable in the experimental group, it was used in the experimental group. Network training was stopped when the prediction error in the training group was minimized using the mean square error. In the next step, the strength of ANN models and the logistic regression were compared with each other. The difference between the values predicted by the two models and the actual values of the output variable in each patient was compared using the mean squared difference method. This statistic follows the non-central Chi-square distribution. In addition, to obtain clinically tangible results, the diagnostic power of the models for separating deceased patients from normal individuals was determined by the ROC curve analysis. For this purpose, the differences of the area under the curve in different models were compared with each other.

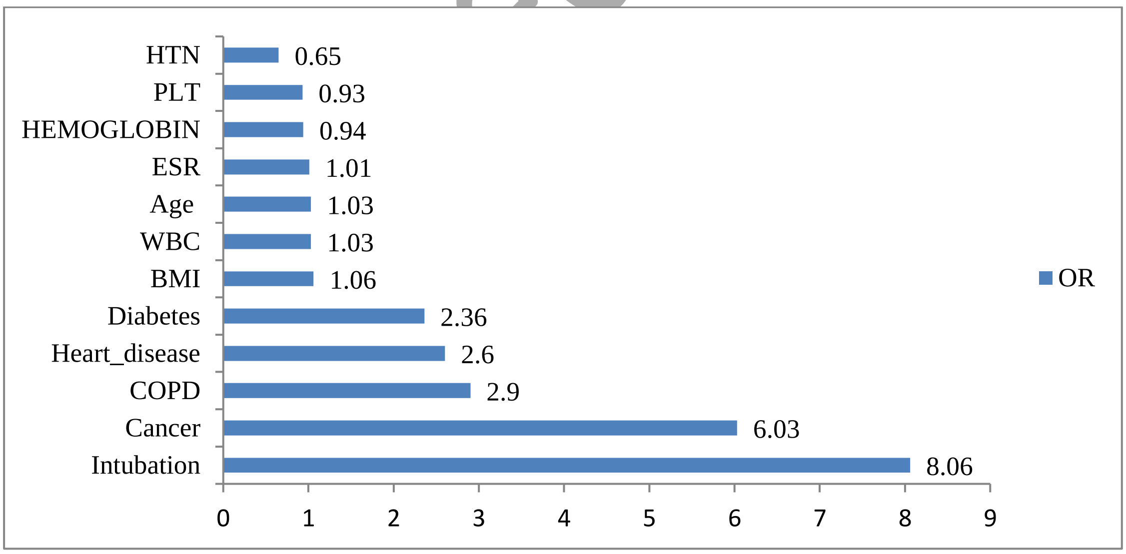
MATLAB and SPSS software were used to teach ANNs and for statistical analysis, respectively.

3. Results

Among the participants, 2206 patients were randomly divided into two training (1544) and experimental (662) groups. The results of the two models in selecting important variables based on their impact factor are shown in Table 1.

Table 1. Status of selected variables in regression model and artificial neural network model

The importance of variables in ANN model	Importance of variables in logistic regression model
Intubation	Intubation
Cancer	COPD
COPD	Heart_disease
Heart_disease	Age
Diabetes	BMI
BMI	Cancer
WBC	Diabetes
Age	ESR
ESR	HEMOGLOBIN
HEMOGLOBIN	HTN
PLT	WBC
HTN	PLT



Graph 1- Values OR calculated by ANN

After ANN analysis between risk factors, a significant difference remained between the 2 groups with regards to intubation (OR 8.06, 95% CI 2.20-5.42, p =0.019), cancer (OR 6.03, 95% CI 1.97-2.45, p =0.046), and COPD (OR 2.9, 95% CI 4. 14-7.12, p = 0.022), heart disease (OR 2.6, 95% CI 3.60-4.18, p = 0.016) and diabetes (OR 2.36, 95% CI 2.55-3.11, p = 0.011).(Graph 1)

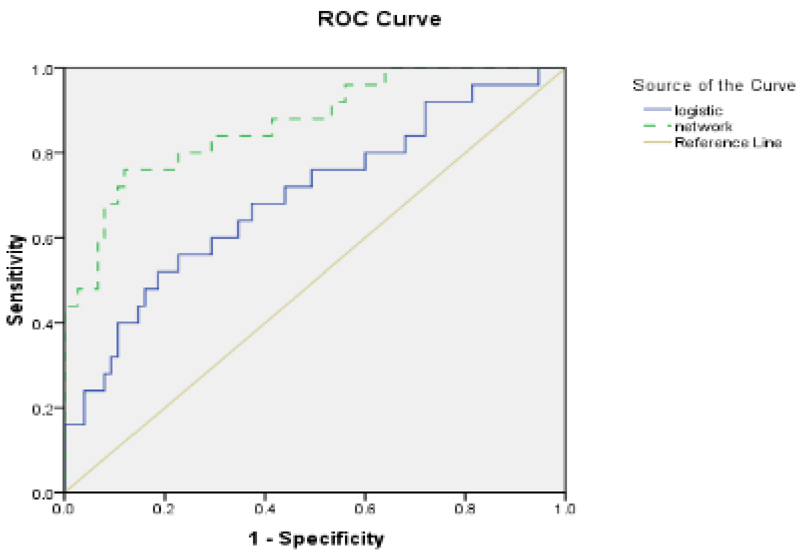


Figure 1- Results of ROC model comparing Logistic Regression (LR) and Neural Network Models (ANNs)

According to Figure 1, the accuracy rate for the best neural network model and logistic regression were 89.76 and 78.66, respectively. The values obtained for this index show the superiority of the ANN model over the logistic regression model.

4. Discussion

In this study, we predicted the variables that had an important role in the mortality of COVID-19 patients. Twelve variables were selected using ANN models. Understanding the relationship between different factors and variables and predicting events, such as mortality, is of interest to many health researchers and policy-makers and can pave the way for better decision-making in this area.

We noticed having intubation, cancer, COPD, heart disease and diabetes has greatly influenced mortality incidence as male patients have been shown to be more vulnerable, which is inconsistent with the results of the study conducted by other studies (22-25).

Also, the results of our study confirmed that an increased age (> 55 years) was associated with death in COVID-19 patients. Moreover, previous evidence confirmed that COVID-19 infection was more likely to affect older males with comorbidities, and could result in fatal respiratory diseases, such as acute respiratory disease syndrome(26). In addition, the increased age may have other risk factors, such as diabetes, respiratory system disease, and cardiovascular disease, which is consistent with other studies(27, 28).

Previously, evidence confirmed that hypertension and diabetes were more prevalent in patients with severe MERS infection. Correspondingly, the mortality rate of influenza was significantly higher in patients with hypertension, metabolic disease, CVDs, and respiratory system disease (29-31).

Intubation, older age (> 50 years old), BMI, hypertension, heart disease, diabetes, COPD, and cancer were associated with a greater risk of death from COVID-19 infection (32). The results of the present ANN analysis could help clinicians to identify high-risk groups that should receive off-label medications or invasive supportive care as soon as possible.

Considering that almost all forecast models existing supplier use linear and logistic methods for the analysis, use the nonlinear relationships discovered in systems ANN can design more effective screening programs people who died from COVID-19. The use of these models, especially in population screening programs estimating small amounts can cause drastic changes in reducing mistakes to show their superiority. This improvement makes screening programs more targeted to find people with a high risk of death, resulting in improved coverage, cost, and effectiveness of these programs.

Ethical considerations

This article is approved by the Medical Ethics Committee of the Ilam University of Medical Science, Ilam, Iran (approval No. IR.MEDILAM.REC.1399.247). The principles of research ethics, honesty, and transparency were considered in all stages of the study.

Funding Acknowledgement

This research project was financially supported by Ilam University of Medical Sciences, Ilam, Iran.

Conflict of interest

The authors declare no conflict of interest.

Contribution:

TF developed and designed the evaluation, collected the clinical data (TF) and drafted the manuscript. TF and ESh participated in the study conception and design, supervised the study and revised the manuscript critically for important intellectual contents. ESh revised the manuscript critically for important intellectual contents. All authors read and approved the final manuscript.

Abbreviations:

ACE2: angiotensin-converting enzyme

COVID-19: coronavirus disease 2019

MERS; Middle East respiratory syndrome

SARS-CoV; severe acute respiratory syndrome coronavirus.

SARS-CoV-2; severe acute respiratory syndrome coronavirus 2.

Reference

1. D'Amico F, Baumgart DC, Danese S, Peyrin-Biroulet L. Diarrhea during COVID-19 infection: pathogenesis, epidemiology, prevention and management. Clinical Gastroenterology and Hepatology. 2020.
2. Brown RJ, Head MG. Monitoring investments in coronavirus research and development. The Lancet Microbe. 2020;1(2):e61.
3. Calina D, Sarkar C, Arsene AL, Salehi B, Docea AO, Mondal M, et al. Recent advances, approaches and challenges in targeting pathways for potential COVID-19 vaccines development. Immunol Res. 2020:1-10.
4. Forni G, Mantovani A. COVID-19 vaccines: where we stand and challenges ahead. Cell Death & Differentiation. 2021;28(2):626-39.
5. Wouters OJ, Shadlen KC, Salcher-Konrad M, Pollard AJ, Larson HJ, Teerawattananon Y, et al. Challenges in ensuring global access to COVID-19 vaccines: production, affordability, allocation, and deployment. The Lancet. 2021;397(10278):1023-34.
6. Mullard A. How COVID vaccines are being divvied up around the world. Nature. 2020.
7. Mirzaee S, Jalalinejad R, George B. Country of Origin, Covid-19 Vaccine and the Future of Travel: a Preliminary Study in Iran. Anais Brasileiros de Estudos Turísticos-ABET. 2021.
8. Nourmohammadi H, Azami G, Taghinezhad F, Shafiei EJCI. Risk and benefit of antiplatelet therapy in Covid-19 patients with cancer and thrombocytopenia: letter to the editor. 2021;39(3):217-8.
9. Mills MC, Salisbury D. The challenges of distributing COVID-19 vaccinations. EClinicalMedicine. 2021;31.
10. Nademi A, Shafiei E, Valizadeh RJIJoP, Sciences B. Forecasting Psychological Symptoms in Mild Traumatic Brain Injury: A Nonparametric Approach. 2019;13(3).

11. Shafiei E, Fakharian E, Omidi A, Akbari H, Delpisheh A, Nademi AJIjop, et al. Comparison of artificial neural network and logistic regression models for prediction of psychological symptom six months after mild traumatic brain injury. 2017;11(3).
12. Niazkar HR, Niazkar M. Application of artificial neural networks to predict the COVID-19 outbreak. Global Health Research and Policy. 2020;5(1):50.
13. Li Y, Xia L. Coronavirus disease 2019 (COVID-19): role of chest CT in diagnosis and management. Am J Roentgenol. 2020;214(6):1280-6.
14. Levi M, Thachil J, Iba T, Levy JH. Coagulation abnormalities and thrombosis in patients with COVID-19. Lancet Haematol. 2020;7(6):e438-e40.
15. Abdulaal A, Patel A, Charani E, Denny S, Mughal N, Moore L. Prognostic modeling of COVID-19 using artificial intelligence in the United Kingdom: model development and validation. J Med Internet Res. 2020;22(8):e20259.
16. Fei Y, Hu J, Gao K, Tu J, Li W-q, Wang W. Predicting risk for portal vein thrombosis in acute pancreatitis patients: A comparison of radical basis function artificial neural network and logistic regression models. Journal of critical care. 2017;39:115-23.
17. Liu Q, Shan M, Zhu H, Cao J, Chen R. Noninvasive ventilation with a helmet in patients with acute respiratory failure caused by chest trauma: a randomized controlled trial. Sci Rep. 2020;10(1):21489.
18. Zhang Z, Beck MW, Winkler DA, Huang B, Sibanda W, Goyal H. Opening the black box of neural networks: methods for interpreting neural network models in clinical applications. Annals of translational medicine. 2018;6(11).
19. Paul PK, Chatterjee D, Ghosh M. Neural Networks: Emphasizing its Application in the World of Health and Medical Sciences. Journal of Advances in Medicine. 2012;1(2):93-9.
20. Moftakhar L, Mozhgan S, Safe MS. Exponentially increasing trend of infected patients with COVID-19 in Iran: a comparison of neural network and ARIMA forecasting models. Iranian Journal of Public Health. 2020.
21. Zheng M, Wu C, Yang K, Yang Y, Liu Y, Gao S, et al. Novel selective hexokinase 2 inhibitor Benitrobenrazide blocks cancer cells growth by targeting glycolysis. Pharmacol Res. 2020:105367.
22. Sepandi M, Taghdir M, Alimohamadi Y, Afrashteh S, Hosamirudsari HJIjoph. Factors associated with mortality in COVID-19 patients: a systematic review and meta-analysis. 2020;49(7):1211.
23. Zhang X-B, Hu L, Ming Q, Wei X-J, Zhang Z-Y, Chen L-D, et al. Risk factors for mortality of coronavirus disease-2019 (COVID-19) patients in two centers of Hubei province, China: A retrospective analysis. 2021;16(1):e0246030.
24. Albitar O, Ballouze R, Ooi JP, Ghadzi SMSJDr, practice c. Risk factors for mortality among COVID-19 patients. 2020;166:108293.
25. Jordan RE, Adab P, Cheng K. Covid-19: risk factors for severe disease and death. British Medical Journal Publishing Group; 2020.
26. Leung CJRimv. Clinical features of deaths in the novel coronavirus epidemic in China. 2020;30(3):e2103.
27. Canning D, Karra M, Dayalu R, Guo M, Bloom DE. The association between age, COVID-19 symptoms, and social distancing behavior in the United States. MedRxiv. 2020.
28. Roncon L, Zuin M, Rigatelli G, Zuliani G. Diabetic patients with COVID-19 infection are at higher risk of ICU admission and poor short-term outcome. Journal of Clinical Virology. 2020;127:104354.
29. Badawi A, Ryoo SGJIJoID. Prevalence of comorbidities in the Middle East respiratory syndrome coronavirus (MERS-CoV): a systematic review and meta-analysis. 2016;49:129-33.
30. Patel SK, Velkoska E, Freeman M, Wai B, Lancefield TF, Burrell LMJFip. From gene to protein—experimental and clinical studies of ACE2 in blood pressure control and arterial hypertension. 2014;5:227.
31. Patel SK, Velkoska E, Burrell LMJC, Pharmacology E, Physiology. Emerging markers in cardiovascular disease: where does angiotensin-converting enzyme 2 fit in? 2013;40(8):551-9.
32. Ryan JJ, Melendres-Groves L, Zamanian RT, Oudiz RJ, Chakinala M, Rosenzweig EB, et al. Care of patients with pulmonary arterial hypertension during the coronavirus (COVID-19) pandemic. Pulmonary circulation. 2020;10(2):2045894020920153.