COVID-19 screening: use of an artificial neural network

Narges Roustaei [ORCID],[1] Elahe Allahyari [ORCID] [2]

[1]Biostatistics and Epidemiology, Yasuj University of Medical Sciences, Yasuj, Iran (the Islamic Republic of)
[2]Department of Epidemiology and Biostatistics, Birjand University of Medical Sciences, Birjand, Iran (the Islamic Republic of)

Correspondence to
Dr Elahe Allahyari, Department of Epidemiology and Biostatistics, School of Health, Birjand University of Medical Sciences, Ghafari Street, Birjand, Iran; elaheh.allahyari@gmail.com

Check for updates

ABSTRACT

Objectives COVID-19 is the biggest pandemic of the 21st century. The disease can be influenced by various sociodemographic factors and can manifest as clinical, pulmonary and gastrointestinal symptoms. This study used an artificial neural network (ANN) model with important sociodemographic factors as well as clinical, pulmonary and gastrointestinal symptoms to screen patients for COVID-19. Patients themselves can screen for these symptoms at home.

Methods Data on all registered patients were extracted in autumn. The best ANN model was selected from different combinations of connections, some hidden layers and some neurons in each hidden layer. In this study, 70% of the data were used in the network training process and the remaining 30% were used to evaluate the function of the multilayer, feed-forward, back-propagation algorithm.

Results The sensitivity and specificity of the ANN model in diagnosing patients with COVID-19 were 94.5% and 17.4%. In order of priority, clinical symptoms, sociodemographic factors, pulmonary symptoms and gastrointestinal symptoms were important predictive factors for COVID-19 using the ANN model. Screening patients for COVID-19 using clinical symptoms and sociodemographic factors (80% importance) remains essential.

Conclusions Home monitoring of oxygen saturation and body temperature as well as old age and drug addiction can be helpful in self-screening symptoms of COVID-19 at home, thereby preventing unnecessary visits to medical centres and reducing burden on medical services.

Key messages

► Most of the COVID-19 symptoms have been identified and assessed separately in previous studies.
► The pattern of the relationship between different factors remains unknown for home screening of the disease.
► Our screening method had high sensitivity and acceptable specificity, with approximately 95% of patients with COVID-19 identified by home monitoring.

INTRODUCTION

COVID-19 was identified from a series of cases of an acute atypical respiratory disease in Wuhan, Hubei Province in the People's Republic of China in December 2019.[1][2] The COVID-19 pandemic has spread around the world, with 4.6 million deaths across almost 200 countries.[2][3] According to WHO, there have been 255 324 963 confirmed cases of COVID-19, including 5 127 696 deaths, as of 17 November 2021. (for up-to-date figures, see https://who.sprinklr.com/).

Various sociodemographic factors and clinical, pulmonary and gastrointestinal symptoms can be seen in patients with COVID-19. Sociodemographic factors such as older age, disease duration, drug addiction and direct contact with patients can be predictive factors for COVID-19. In recent epidemiological studies, old age (usually older than 60 years old) has been indicated to be a high-risk factor for contracting severe COVID-19.[2–4] The risk of severe COVID-19 and death can also escalate with drug addiction as it can cause inflammation and reduce the lungs' ability to respond to infection.[5][6] Moreover, this type of virus can spread by inhaling droplets of saliva from sneezing and coughing, or even by exhaling or touching droplets on hands and face, and therefore direct contact with patients and lack of social distancing can effectively spread the virus.

Clinical symptoms such as saturation, body temperature, dizziness, headache, myalgia, loss of consciousness, hyposmia, paresis, weakness and tremor could also affect the diagnosis of COVID-19.[7][8] Pulmonary symptoms such as respiratory rate, cough, sore throat, chest pain, dyspnoea and respiratory distress are common symptoms that appear 2–14 days after exposure to the virus.[9] Nausea,

BMJ

Table 1 Weight of trained artificial neural network for PCR response

		First hidden layer					Second hidden layer		
		Node 1	Node 2	Node 3	Node 4	Node 5	Node 1	Node 2	Node 3
Input layer	Drug addiction								
	No	0.12	−0.822	−0.414	0.053	−0.503			
	Yes	−0.8	0.32	−0.055	0.211	0.241			
	Respiratory rate								
	5–10	−0.271	−0.133	0.228	0.143	−0.343			
	10–14	0.02	0.42	0.412	−0.119	0.216			
	14–18	−0.548	−0.173	−0.032	0.149	−0.408			
	18–22	0.087	−0.309	0.25	0.258	0.297			
	22–28	0.153	−0.241	−0.289	−0.049	0.288			
	≥28	−0.212	0.291	0.026	−0.251	−0.203			
	Contact								
	No	−0.002	−0.115	0.361	−0.36	0.517			
	Yes	0.345	−0.247	−0.169	−0.277	−0.264			
	Cough								
	No	0.032	0.428	−0.197	0.471	0.408			
	Yes	−0.251	0.271	0.325	0.232	0.064			
	Sore throat								
	No	0.436	−0.439	−0.182	−0.319	0.397			
	Yes	0.373	−0.234	0.264	0.399	−0.505			
	Chest pain								
	No	0.442	−0.433	−0.445	−0.357	−0.141			
	Yes	0.356	−0.39	0.192	0.212	−0.088			
	Dyspnoea								
	No	−0.314	0.21	0.216	−0.474	0.319			
	Yes	0.146	−0.248	−0.225	0.432	0.222			
	Respiratory distress								
	No	−0.625	0.115	0.562	−0.367	0.003			
	Yes	0.444	−0.761	−0.018	−0.594	−0.313			
	Dizziness								
	No	−0.303	−0.019	0.077	0.11	0.071			
	Yes	0.001	−0.05	−0.41	−0.076	0.427			
	Headache								
	No	−0.046	−0.144	0.358	−0.358	−0.128			
	Yes	−0.369	−0.19	−0.196	0.291	0.021			
	Myalgia								
	No	−0.297	0.316	0.164	0.435	−0.044			
	Yes	0.299	−0.254	0.217	0.118	−0.374			
	Reduced consciousness								
	No	0.448	−0.179	−0.39	−0.343	−0.259			
	Yes	−0.133	0.228	−0.054	0.222	0.027			
	Hyposmia								
	No	0.258	0.024	0.298	−0.455	−0.396			
	Yes	0.34	0.075	0.338	−0.33	−0.467			
	Paresis								
	No	0.127	0.348	0.072	−0.117	0.16			
	Yes	0.004	0.395	0.278	0.396	0.152			
	Weakness								
	No	0.641	−0.263	0.25	0.272	0.377			
	Yes	0.033	−0.259	−0.257	−0.438	−0.031			
	Tremor								
	No	0.183	−0.327	−0.21	−0.141	−0.123			

Continued

Table 1 Continued

		First hidden layer					Second hidden layer		
		Node 1	Node 2	Node 3	Node 4	Node 5	Node 1	Node 2	Node 3
	Yes	0.17	0.507	0.024	0.377	0.089			
	Nausea								
	No	−0.184	0.342	0.086	0.413	0.304			
	Yes	−0.092	0.3	−0.223	−0.194	−0.336			
	Vomiting								
	No	0.183	−0.075	0.251	0.142	0.252			
	Yes	0.207	0.284	0.042	0.422	0.102			
	Diarrhoea								
	No	−0.262	−0.103	0.465	−0.103	0.027			
	Yes	0.505	0.169	−0.38	−0.124	−0.356			
	Anorexia								
	No	−0.271	0.323	−0.007	−0.019	−0.228			
	Yes	0.101	0.197	−0.176	0.199	−0.216			
	Age	0.244	−0.193	0.09	−0.154	−0.18			
	Duration	10.058	−0.782	−0.18	−0.86	−0.06			
	Saturation by pulse oximetry	−0.678	0.45	−0.144	−0.019	0.452			
	Temperature	−0.439	0.36	−0.039	0.085	−0.086			
Second hidden layer	Node 1	0.306	0.111	0.151	−0.274	0.055			
	Node 2	−1.057	0.492	−0.427	−0.053	1.392			
	Node 3	0.215	−0.106	−0.21	−0.017	0.007			
Output layer	PCR								
	Negative						0.707	0.963	−0.054
	Positive						−0.721	−0.792	0.317

vomiting, diarrhoea and anorexia are known gastrointestinal symptoms that are also significant in screening for COVID-19.[10]

Purpose

The lack of efficient screening methods has created many problems and limitations all over the world. Furthermore, the high transmission rate of COVID-19 revealed a deficiency of applicant screening systems more and more in this regard. Therefore, COVID-19 is affecting more and more people each day, along with the ability of the virus to mutate into various types. Since the pattern of the relationship between different factors remains unknown in this disease, it is necessary to use methods such as an artificial neural network (ANN), which determines the pattern between variables, while not having restrictive underlying assumptions such as normality. In this study, the ANN model was used to screen patients for COVID-19 using some critical sociodemographic factors and clinical, pulmonary and gastrointestinal symptoms, which can be screened by people themselves at home.

METHODS

Most of the symptoms in the medical care monitoring centres were registered for suspected COVID-19 patients who had PCR tests. Data on all registered patients in South Khorasan Province were extracted from 20 May to 20 December 2020; this allowed us to assess common disease symptoms in spring, summer and autumn separately. After removing patients with missing data, PCR test results were available only for 29 (7 cases tested negative and 22 tested positive), 24 (12 cases tested negative and 12 tested positive) and 1612 (453 cases tested negative and 1159 tested positive) participants in spring, summer and autumn, respectively. The small sample size in spring and summer limited our study to people who had PCR tests in autumn. Data collected in a checklist included age, oxygen saturation (SpO_2) by pulse oximetry, body temperature, disease duration, cigarette or drug addiction (0 no, 1 yes), respiratory rate (5–10, 10–14, 14–18, 18–22, 22–28 and >28), direct contact with the patient (0 no, 1 yes), and having any symptoms including cough, sore throat, chest pain, dyspnoea, respiratory distress, dizziness, headache, myalgia, loss of consciousness, hyposmia, paresis, weakness, tremor, nausea, vomiting, diarrhoea and anorexia, and a PCR test result (0 negative, 1 positive). The PCR test result (0 negative, 1 positive) was a response variable.

Data analysis

Multilayer, feed-forward, back-propagation, supervised neural networks are powerful neural networks used in prediction and modelling construction.[11] We used these algorithms, and accordingly the best
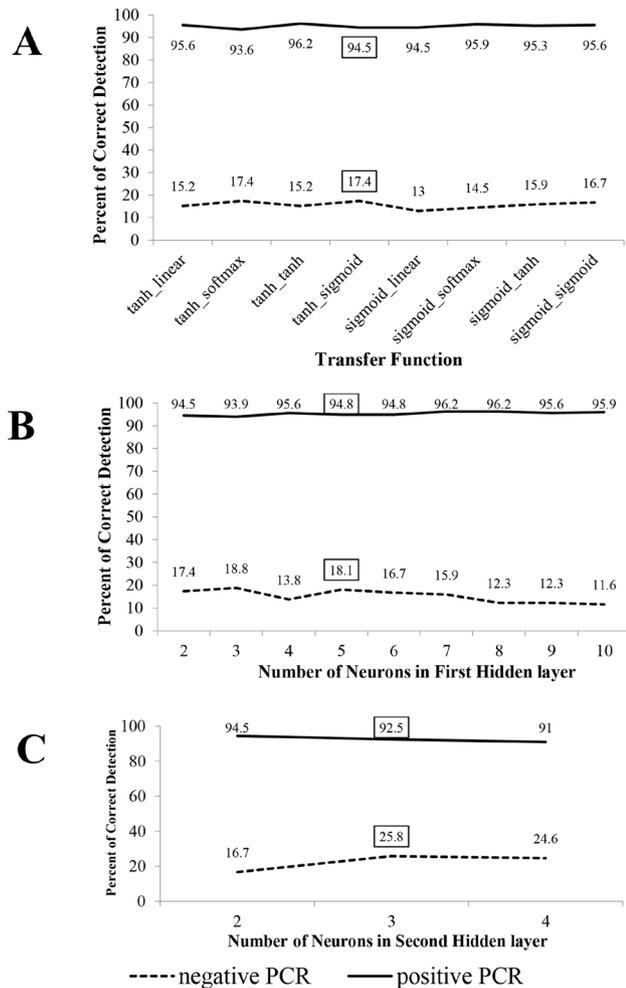
Figure 1 (A) Per cent of correct detection of the ANN model for different transfer functions. (B) Per cent of correct detection of the ANN model for different number of neurons in the first hidden layer. (C) Per cent of correct detection of the ANN model for different number of neurons in the second hidden layer. ANN, artificial neural network.

structure of the neural network was selected from different combinations of connection functions (hyperbolic tangent or sigmoid to connect the input and middle layer neurons, linear, hyperbolic tangent, or sigmoid functions to connect the middle and output layer neurons), number of hidden layers (one or two) and number of neurons in each hidden layer (between 2 and 10).[11] Of the data, 70% were used in the network training process and the remaining were applied in model evaluation. We repeated each grid three times to avoid random assignment of weights and random correlations. We report the best results. Lastly, the importance of input variables in the diagnosis of patients was determined. All analyses were performed using the Statistical 100 Package for Social Sciences V.26.

RESULTS

Data on 1612 subjects (720 women, 892 men) were assessed. Only 10 subjects were the healthy staff. The mean±SD age of the participants was 58.83±19.84 years. Of the patients who had positive PCR, 115 died and 158 were hospitalised.

Figure 1A shows that the neural network with hyperbolic tangent and sigmoid functions that connect the hidden-input and hidden-output layers with two neurons in the middle layer had acceptable sensitivity (94.5%) and the highest specificity (17.4%). Increasing the number of neurons in the hidden layer and increasing the number of hidden layers improved the performance of the model (figure 1B, C). Therefore, the best ANN model had two hidden layers with five and three neurons and with a hyperbolic tangent and sigmoid functions that connect the hidden-input, hidden-hidden and hidden-output layers, respectively.

Clinical and sociodemographic factors had a nearly 80% importance in the prediction, whereas gastrointestinal and pulmonary symptoms showed minor importance in the prediction with 6% and 14.5%. Clinical symptoms such as body temperature and SpO_2, as well as sociodemographic factors such as duration, age and drug addiction, played an essential role in the diagnosis of COVID-19. On the other hand, pulmonary symptoms such as respiratory distress, sore throat and respiratory rate, as well as gastrointestinal symptoms such as nausea and diarrhoea, had the highest importance.

Table 1 shows that respiratory rate lower than 18, direct contact with the patient, and symptoms such as sore throat, chest pain, dyspnoea, respiratory distress, myalgia, hyposmia, weakness, nausea, diarrhoea, anorexia, decreasing SpO_2 and body temperature, as well as ageing and disease duration, increase the likelihood of diagnosis of COVID-19.

DISCUSSION

In order of priority, clinical symptoms, sociodemographic factors, pulmonary symptoms and gastrointestinal symptoms were important predictive factors for COVID-19. In addition, screening patients for COVID-19 using clinical and sociodemographic factors (80% importance) remains essential. Among the different clinical symptoms, body temperature and SpO_2 were the most important predictors that can be measured at home by pulse oximetry and thermometer without visiting medical centres. Similar studies have reported that an SpO_2 of less than 92 and a body temperature greater than 37.5°C were clinical factors that can identify patients with COVID-19.[7 8]

Age, drug addiction and disease duration were important sociodemographic factors that predict COVID-19. Studies have shown that susceptibility to this infection increases with increasing age.[4] Overall, greater attention must be given to older people and they should be trained to observe social distancing and healthcare measures. The more time that passed from the infection, the more likely it is to detect patients with COVID-19. For people with addiction, having

access to healthcare services can be more complex and so they attempt to obtain drugs in unhealthy ways.[12] People with addictive disorders need easier accessibility to prescription drugs and healthcare services.[12]

In this study, a respiratory rate lower than 18 indicated pulmonary symptoms among people with COVID-19. Some studies have indicated that people with COVID-19 also had sore throats.[9] Also, most patients with COVID-19 experience respiratory distress in the severe stage, leading to hospitalisation and intubation and admission to the intensive care unit.[10] Patients may also experience acute gastrointestinal symptoms including nausea and diarrhoea. In a systematic review and meta-analysis, 15.47% of patients had at least one gastrointestinal symptom.[13]

The sensitivity and specificity of the ANN model in this study were 94.5% and 17.4%, respectively. In other words, about 95% of patients with COVID-19 were identified by home monitoring using this ANN model. A meta-analysis showed CT scan presented a pooled sensitivity of 95.48%.[14] Therefore, the sensitivity of this ANN model and the CT scan in identifying patients with COVID-19 was approximately the same.

COVID-19 transmission and mutation rates are very high[15] and early screening for COVID-19 is essential. Therefore, the advantage of having a high sensitivity is preventing the wider spread of the disease in the community, resulting in reduced cost of treatment, reduced work pressure among the medical staff in hospitals and medical centres, and reduced economic, social, cultural and political problems due to COVID-19.

CONCLUSION

ANN is a relatively novel technique used to predict complex non-linear relationships in epidemiology and public health. Home monitoring of SpO_2 and body temperature as well as old age and drug addiction can be helpful when self-screening for COVID-19 at home, with the ANN having an accuracy of 94.5%.

ORCID iDs
Narges Roustaei http://orcid.org/0000-0002-5753-2142
Elahe Allahyari http://orcid.org/0000-0003-3859-4820

REFERENCES

1. Cao X. COVID-19: immunopathology and its implications for therapy. *Nat Rev Immunol* 2020;20:269–70.
2. Yuki K, Fujiogi M, Koutsogiannaki S. COVID-19 pathophysiology: a review. *Clin Immunol* 2020;215:108427.
3. Velavan TP, Meyer CG. The COVID-19 epidemic. *Trop Med Int Health* 2020;25:278–80.
4. Davies NG, Klepac P, Liu Y, et al. Age-Dependent effects in the transmission and control of COVID-19 epidemics. *Nat Med* 2020;26:1205–11.
5. Volkow ND. *Collision of the COVID-19 and addiction epidemics.*. American College of Physicians, 2020: 173. 61–2.
6. Marsden J, Darke S, Hall W, et al. Mitigating and learning from the impact of COVID-19 infection on addictive disorders. *Addiction* 2020;115:1007–10.
7. Shah S, Majmudar K, Stein A, et al. Novel use of home pulse oximetry monitoring in COVID-19 patients discharged from the emergency department identifies need for hospitalization. *Acad Emerg Med* 2020;27:681–92.
8. Torjesen I. *Covid-19: patients to use pulse oximetry at home to spot deterioration.* British Medical Journal Publishing Group, 2020.
9. Weng L-M, Su X, Wang X-Q. Pain symptoms in patients with coronavirus disease (COVID-19): a literature review. *J Pain Res* 2021;14:147–59.
10. Duong L, Xu P, Liu A. *Meningoencephalitis without respiratory failure in a young female patient with COVID-19 infection in Downtown Los Angeles, early* April 2020. *Brain Behav Immun* 2020;87:33.
11. Ghaedi AM, Vafaei A. Applications of artificial neural networks for adsorption removal of dyes from aqueous solution: a review. *Adv Colloid Interface Sci* 2017;245:20–39.
12. Dubey MJ, Ghosh R, Chatterjee S, et al. COVID-19 and addiction. *Diabetes Metab Syndr* 2020;14:817–23.
13. Ghimire S, Sharma S, Patel A, et al. Diarrhea is associated with increased severity of disease in COVID-19: systemic review and metaanalysis. *SN Compr Clin Med* 2021;28–35.
14. Esposito A, Palmisano A, Scotti GM. Why is chest CT important for early diagnosis of COVID-19? prevalence matters. *medRxiv* 2020.
15. Mercer TR, Salit M. Testing at scale during the COVID-19 pandemic. *Nat Rev Genet* 2021;22:415–26.