



Application of artificial intelligence in COVID-19 medical area: a systematic review

Zhoulin Chang^{1#}, Zhiqing Zhan^{2#}, Zifan Zhao³, Zhixuan You³, Yang Liu⁴, Zhihong Yan⁵, Yong Fu⁵, Wenhua Liang⁶, Lei Zhao⁷

¹College of Mechanical and Electrical Engineering, Guangdong University of Science and Technology, Dongguan, China; ²The Third Clinical College, Guangzhou Medical University, Guangzhou, China; ³Nanshan College, Guangzhou Medical University, Guangzhou, China; ⁴School of Information Engineering, Zhengzhou University, Zhengzhou, China; ⁵Kuangji Medical Technology (Guangdong Hengqin) Co., Ltd., Zhuhai, China; ⁶Department of Thoracic Surgery, The First Affiliated Hospital of Guangzhou Medical University, Guangzhou, China; ⁷Department of Physiology, School of Basic Medical Sciences, Guangzhou Medical University, Guangzhou, China

Contributions: (I) Conception and design: W Liang, L Zhao, Z Chang; (II) Administrative support: Z Chang; (III) Provision of study materials or patients: All authors; (IV) Collection and assembly of data: Z Zhan, Z Zhao, Z You; (V) Data analysis and interpretation: All authors; (VI) Manuscript writing: All authors; (VII) Final approval of manuscript: All authors.

[#]These authors contributed equally to this work.

Correspondence to: Wenhua Liang. Department of Thoracic Surgery, The First Affiliated Hospital of Guangzhou Medical University, Guangzhou 510120, China. Email: liangwh1987@163.com; Lei Zhao. Department of Physiology, School of Basic Medical Sciences, Guangzhou Medical University, Guangzhou 511436, China. Email: crystal-zl@163.com.

Background: Coronavirus disease 2019 (COVID-19) has caused a large-scale global epidemic, impacting international politics and the economy. At present, there is no particularly effective medicine and treatment plan. Therefore, it is urgent and significant to find new technologies to diagnose early, isolate early, and treat early. Multimodal data driven artificial intelligence (AI) can potentially be the option. During the COVID-19 Pandemic, AI provided cutting-edge applications in disease, medicine, treatment, and target recognition. This paper reviewed the literature on the intersection of AI and medicine to analyze and compare different AI model applications in the COVID-19 Pandemic, evaluate their effectiveness, show their advantages and differences, and introduce the main models and their characteristics.

Methods: We searched PubMed, arXiv, medRxiv, and Google Scholar through February 2020 to identify studies on AI applications in the medical areas for the COVID-19 Pandemic.

Results: We summarize the main AI applications in six areas: (I) epidemiology, (II) diagnosis, (III) progression, (IV) treatment, (V) psychological health impact, and (VI) data security. The ongoing development in AI has significantly improved prediction, contact tracing, screening, diagnosis, treatment, medication, and vaccine development for the COVID-19 Pandemic and reducing human intervention in medical practice.

Discussion: This paper provides strong advice for using AI-based auxiliary tools for related applications of human diseases. We also discuss the clinicians' role in the further development of AI. They and AI researchers can integrate AI technology with current clinical processes and information systems into applications. In the future, AI personnel and medical workers will further cooperate closely.

Keywords: Artificial intelligence (AI); coronavirus disease 2019 (COVID-19); severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2); deep learning; machine learning

Submitted Apr 29, 2021. Accepted for publication Sep 02, 2021.

doi: 10.21037/jtd-21-747

View this article at: <https://dx.doi.org/10.21037/jtd-21-747>

Introduction

The coronavirus disease 2019 (COVID-19) pandemic has caused extreme strains on health systems, public health infrastructure, and many countries' economies. It was caused by the new severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) (1,2). When talking, coughing, or sneezing, droplets sprayed from the COVID-19 patients were the most common coronavirus transmission mode. Surface contact transmission (touching an object's surface with the virus) was another possible mode (3). The clinical symptoms of COVID-19 patients are fever, cough, shortness of breath, chills, trembling, muscle pain, headache, sore throat, loss of taste or smell, etc. Severe patients often develop dyspnea or hypoxemia 1 week after the onset. In severe cases, they rapidly progress to acute respiratory distress syndrome (ARDS), septic shock, metabolic acidosis, coagulation dysfunction, and multiple organ failure (4,5).

Globally, as of March 2021, there have been 119,220,681 confirmed cases of COVID-19, including 2,642,826 deaths, reported to the World Health Organization. The number of affected countries, areas, or territories was 223. More than billions of people were staying home to avoid SARS-CoV-2. At the same time, a large number of problems emerged (6). For example, the number of hospital beds and doctors was scarce. Protective equipment was lacking, and there were no specific drugs. How to disinfect to protect medical staff was also essential. Most countries have implemented a variety of infection control measures, and hospitals have adopted various treatment options. It is currently difficult to determine which steps and programs are better. Since there is no effective treatment, the best way to deal with the SARS-CoV-2 infection is to control the source of infection, diagnose, report, isolate, support treatment, and release epidemic information in time to avoid unnecessary panic. There is an urgent need to explore a high-efficient way to assist human experts in overcoming the COVID-19 Pandemic. Multimodal data drove artificial intelligence (AI) could potentially be the option.

AI technology is gradually moving out of the laboratory toward clinical and public health applications in recent years, such as early warning of epidemics and intelligent analysis of extensive medical data. In fighting on the COVID-19, AI had dramatically improved our diagnosis, prediction, and treatment level (7-9). AI could analyze the epidemiological characteristics, clinical characteristics, and treatment effects of COVID-19 through extensive data of clinical cases. AI was also used for quantitative digital

analysis of medical images and guide diagnosis. It solidified the expert's knowledge system into the model. As the number of learning samples increased, the accuracy of the analysis increased. So, it could guide the identification and treatment of the COVID-19 patients. Scientists had made various new COVID-19 prediction models based on big data, providing an essential basis for policy formulation. Through the big data of the COVID-19 patients, many countries have known about the incubation period and clinical characteristics. This article provides more details for AI applying in human diseases.

We present the following article in accordance with the PRISMA reporting checklist (available at <https://dx.doi.org/10.21037/jtd-21-747>).

Methods

We searched PubMed, arXiv, medRxiv, and Google Scholar through February 2020 to identify studies on AI applications in the medical areas for the COVID-19 Pandemic. We used the following search terms: [(“artificial intelligence” OR “AI”) OR (“machine learning” OR “ML”) OR (“deep learning” OR “DL”)] AND (“covid-19” OR “sars-cov-2” OR “Coronavirus” OR “pandemic”) (Table S1). The initial literature search identified 144,735 articles. After removal of duplicates, an initial screen was conducted to ensure that articles were related to AI applications for the COVID-19 Pandemic in the following six medical areas: (I) epidemiology, (II) diagnosis, (III) progression, (IV) treatment, (V) psychological health impact, and (VI) data security. The title of each article was reviewed for topic relevance, and the abstract was reviewed for further clarification as necessary. Eighty-two articles were finally included in this narrative review.

Applications of AI in COVID-19 epidemiology

At the beginning of the Pandemic, the governments and people's concerns are concentrated in the following areas: the duration and peak of the COVID-19 Pandemic, the number of infected people, the related influencing factors, and the measurement to reduce infection, etc. A community of applied mathematicians, virologists, epidemiologists, and AI researchers did many works to help the governments making suitable policies in the fighting (2,3,8,9). This section focuses on AI applications in different epidemiological areas, such as establishing models to predict the COVID-19 spread and evaluating the severity,

identifying and tracing the infection case to control the infectious rate, and studying the related influencing factors to prevent the spread (Table 1).

COVID-19 spread prediction and the severity evaluation

The epidemic spread model of infectious diseases is a traditional epidemiological and mathematical problem with crucial practical value. The susceptible-infectious-recovered (SIR) disease model uses the infection data of the disease itself (including the existing confirmed cases, mortality, and the route of infection) to predict the infection speed of a disease in a short time in a particular area (19). The rapid development of AI technology and big data has brought new changes to these traditional infectious disease models. By analyzing unstructured data (such as social media, transportation, news, government statements, etc.), AI can combine these data with traditional infectious disease data to achieve the role of predicting and monitoring large-scale contagious diseases. Dandekar *et al.* (20) encoded the quarantine policy as a vital function in the optimized neural network-augmented SIR model to indicate the outbreak size in Wuhan, China. This model made up for the traditional SIR model's inability to discover the effects of more granular interactions such as social distancing and quarantine policies. Utilizing the susceptible-infected-recovered-dead (SIRD) pandemic model for the United States, Russia, China, and the Syrian Arab Republic, Al-Raei *et al.* (21) found the coefficient values of the COVID-19 infection, recovery, and mortality. It applied the method for the other countries with the COVID-19 Pandemic and determined the Pandemic reproduction rates. Aslan *et al.* (10) used the cumulative number of infected death cases in Hubei, China, to train the susceptible-exposed-susceptible in quarantine-infected (asymptomatic or having mild symptoms)-reported cases-recovered (SEIQR) deterministic type model, highlighting the effects of quarantine and social distancing in Hubei. Then they analyzed the COVID-19 outbreak and predicted the impact of preventive measures and diagnostic tests in Turkey. It developed accurate local prediction tools, which could be coupled to create global models. Yang *et al.* (11) integrated population migration data and the most updated COVID-19 epidemiological data into the susceptible-exposed-infectious-removed (SEIR) model to derive the epidemic curve. Their dynamic SEIR model was effective in predicting the COVID-19 epidemic peaks and sizes. They also used the long short-term memory model (LSTM)

trained on the 2003 SARS dataset, combined with COVID-19-related feature data (transmission rate, mortality, etc.) to predict the spread of COVID-19. These results indicated that the implementation of control measures was valuable in reducing the eventual COVID-19 epidemic size.

LSTM is a standard recurrent neural network (RNN) often used to process data with time series (22). Chimmula *et al.* (12) developed LSTM networks to predict Canada's pandemic trend and ending point. It captured the transmission dynamics with minimum loss. The root mean squared errors (RMSE) was 34.83, with an accuracy of 93.4% for short-term predictions in Canada. Meanwhile, the RMSE was about 45.70, with an accuracy of 92.67% for long-term predictions. Kolozsvari *et al.* (13) implemented an RNN to predict the epidemic curve. They concluded that repeated peaks were anticipated on the COVID-19 epidemic curve. It used official databases trained on the currently available data, which were validated by root mean squared logarithmic errors (RMSLE) calculation. The mean of RMSLE was different in different countries. For example, the mean of RMSLE in Hungary was 0.06. In addition to RNN, convolutional neural networks (CNN) can also predict the spread of infectious diseases. By converting one-dimensional data into a two-dimensional matrix and inputting it into a CNN, the cumulative number of confirmed cases in a region can be accurately predicted (14,15). Hu *et al.* (14) proposed a modified stacked auto-encoder (MAE) to model COVID-19's transmission dynamics in China. It employed latent variables to divide 34 provinces or cities into nine groups with the similar transmission. The error rate can be as low as 0.73%. The COVID-19 data samples were lacking initially, so Fong *et al.* (15) developed a type of CNN called polynomial neural network with corrective feedback (PNN + cf) to predict the outbreak. It allowed additional input variables during polynomials formation to augment the little existing data and possessed great superiority with a relatively low error.

Trajectory tracking and infectious rate control

Early case identification, quarantining, and preventing exposure to the communities were crucial pillars in managing COVID-19 Pandemic (23). Based on many data such as travel information, social information, consumption information, and exposure history, AI could establish the network of potential infections. Combining the time when the infected person was diagnosed and the spatial location information of their close contacts, it determined the time

Table 1 Applications of AI in COVID-19 epidemiology

First author [year] (reference)	Country (region)	Modality	Model	Data source	Application area	Result
Aslan <i>et al.</i> [2020] (10)	China/Turkey	Demographic data	SEIQR model	Training dataset: the cumulative number of infected death cases in Hubei, China from January 20, 2020 to March 23, 2020; testing dataset: the cumulative number of infected deaths in Turkey from March 10, 2020 to April 10, 2020	COVID-19 spread prediction and the severity evaluation	Develop accurate local prediction models
Yang <i>et al.</i> [2020] (11)	China	Demographic data	SEIR model/RNN using LSTM	Training dataset: SARS epidemic data between April and June 2003 across the whole of China retrieved from an archived news-site (SOHU); testing dataset: migration and epidemiological data before and after January 23 2020	COVID-19 spread prediction and the severity evaluation	Found control measures to reduce the eventual COVID-19 epidemic size
Chimmula <i>et al.</i> [2020] (12)	Canada	Demographic data	RNN using LSTM	Training dataset: 80% number of confirmed cases, fatalities and recovered patients until March 31, 2020 provided by Johns Hopkins University and Canadian Health authority; testing dataset: 20% number of confirmed cases, fatalities and recovered patients until March 31, 2020 provided by Johns Hopkins University and Canadian Health authority	COVID-19 spread prediction and the severity evaluation	Short term predictions: RMSE error 34.83/accuracy 93.4%; long term predictions: RMSE error 45.70/accuracy 92.67%
Kolozsvari <i>et al.</i> [2021] (13)	Globe	Demographic data	RNN using LSTM	The publicly available datasets from the WHO and Johns Hopkins University for the following countries: Austria, Belgium, China (Hubei), Czech Republic, France, Germany, Hungary, Iran, Italy, Netherlands, Norway, Portugal, Slovenia, Spain, Switzerland, UK and the USA until April 10, 2020	COVID-19 spread prediction and the severity evaluation	Mean of RMSLE of prediction 1: Hungary 0.06/UK 0.234/Italy 0.114/Spain 0.266/Germany 0.147/France 0.513/USA 0.216; mean of RMSLE of prediction 2: Hungary 0.107/UK 0.455/Italy 0.155/Spain 0.181/Germany 0.108/France 0.307/USA 0.528
Hu <i>et al.</i> [2021] (14)	Mainland China/ three other regions (Hong Kong, Macau and Taiwan) in China	Demographic data	CNN/MAE	The total numbers of the accumulated and new confirmed cases in all of China from January 11 to February 27, 2020	COVID-19 spread prediction and the severity evaluation	Error rate: 0.73%
Fong <i>et al.</i> [2020] (15)	China	Demographic data	CNN/PNN + cf	COVID-19 case in China between 21 Jan to 3 Feb 2020	COVID-19 spread prediction and the severity evaluation	Offering the highest possible level of prediction accuracy under the constraints of low data availability and knowledge

Table 1 (continued)

Table 1 (continued)

First author [year] (reference)	Country (region)	Modality	Model	Data source	Application area	Result
Mao et al. [2021] (16)	China	Graph data/ demographic data	Graph database model	(I) Resident trajectory information database from Hubei Province(Health Committee); (II) confirmed Person Information Database (Public Security Department; Health Committee); (III) information database of high-risk groups and close contacts (Epidemic Prevention and Control Headquarters); (IV) hospital fever information database (Epidemic Prevention and Control Headquarters); (V) mobile phone signaling database (communications operator branch) from January 21 to February 9, 2020	Trajectory tracking and infectious rate control	Identify closest contacts and a number of public places with high risk of infection
Pramanik et al. [2020] (17)	Globe	Climatic data/ demographic data	BRT model	The average monthly temperature and the average relative humidity data for January to April 2020 in 230 cities; the maximum number of recorded cases for January to April 2020 in 230 cities	Uncovering climatic factors of COVID-19 spread	AUC: 0.8675
Pasayat et al. [2020] (18)	India	Demographic data	EG model/LR model	The publicly available dataset of India from Humanitarian	Uncovering social factors of COVID-19 spread	Accuracy (the exponential growth model): 90.78%; accuracy (the linear regression model): 99.88%

AI, artificial intelligence; COVID-19, coronavirus disease 2019; SEIQR model, susceptible-exposed-susceptible in quarantine (isolated class)-infected (asymptomatic or having mild symptoms)-reported (infected) cases (hospitalized if get severe symptoms or quarantined if get mild symptoms)-recovered; SEIR, susceptible-exposed-infectious-removed; RNN, recurrent neural network; LSTM, long short-term memory model; CNN, convolutional neural networks; MAE, modified auto-encoders; PNN + cf, polynomial neural network with corrective feedback; BRT model, boosted regression tree model; EG model, exponential growth model; LR model, linear regression model; SARS, severe acute respiratory syndrome; RMSE error, root mean squared errors; RMSLE, root mean squared logarithmic errors; AUC, area under the curve.

point and specific transmission path of possible cross-infection. It could help the disease control departments find possible transmission sources and potential super-infectors and quickly cut off the spread. Mao *et al.* (16) reported how Hainan Province in China relied on the government's big data public service platform to gather epidemic multisource big data and used the graph database model to determine and trace contacts, which achieved good results. It indicated that using a centralized model for digital contact tracing was more effective than the decentralized model in China. Srinivasa Rao *et al.* (24) proposed a mobile phone-based web survey to collect a basic travel history and more common signs and symptoms. AI framework processed the data to evaluate individuals and stratify them into no risk, minimal risk, moderate risk, and high-risk groups. So, it identified COVID-19 cases more quickly and checked susceptibility spread. Another machine learning model was based on Taiwan's national health insurance, customs, and immigration database (25). According to several factors (e.g., travel history), the population was stratified into lower risk or higher risk. Those with higher risk were quarantined at home and were tracked through their mobile phones. Finally, the number of infected patients was far fewer than what was initially anticipated.

Uncovering climatic/geographic/social factors of COVID-19 spread

The COVID-19 Pandemic was showing climatic and geographic patterns in its spread and development. Pramanik *et al.* (17) applied a boosted regression tree (BRT) model to find the climatic influence on COVID-19 transmission risks in 228 cities globally across three climatic zones. The number of positive cases decreased sharply above an average temperature of 10 °C in France, Turkey, the US, the UK, and Germany. Among the tropical countries, COVID-19 in Indian cities is most affected by mean daily temperature and those in Brazil by temperature seasonality. Travaglio *et al.* (26) applied the generalized linear models to find that a slight rise in major fossil fuel-related air pollutants led to a considerable change in England's COVID-19 infectivity and mortality rate. It also showed that a 1 m³ increase of the PM_{2.5} long-term average was associated with a 12% increase in COVID-19 cases. Mishra *et al.* (27) implemented a Köppen-Geiger (KG) climate classification model to analyze the potential relations with climate, geographical location, and the COVID-19 Pandemic worldwide. It demonstrated that

countries in high latitudes with temperate or continental climates were the most vulnerable to this outbreak.

Different countries and governments have taken various measures to deal with the COVID-19 Pandemic. Differences in lockdown, quarantine, and social distancing may also contribute to differences in the severity of the COVID-19 Pandemic. Pasayat *et al.* (18) used the exponential growth model and linear regression model to predict the COVID-19 cases with the continuous flow of the lockdown in India. The accuracy of the exponential growth model and the linear regression model was 90.78% and 99.88%. It showed that lockdown with specific restrictions had a vital role in preventing the COVID-19 from spreading in this current situation. Zhang *et al.* (28) applied supervised learning techniques to identify and train the non-parametric network-based SIR (NP-Net-SIR) model. The trained model analyzed the connection between population flow and cross-regional infection strength. The model's accuracy was 99.6%. They found that the non-lock-down-typed measures reached the same containment consequence as the lock-down. Vaid *et al.* (29) used an AI framework based on policy interventions' timeline. This AI model was based on three approaches—the Bayesian SIR model, Kalman filter, and machine learning. The study showed that dropped in the COVID-19 infections' effective growth rate was sharper in stringent policies (the USA and Canada) but was more gradual in the relaxed approach (Sweden).

Applications of AI in COVID-19 diagnosis

At present, the COVID-19 clinical diagnosis is mainly based on different information such as epidemiological history, clinical symptoms, laboratory examination results, chest imaging findings, nucleic acid detection (pharyngeal swabs or serum antibodies), or homologous comparison of gene sequencing. Different methods have different advantages and disadvantages. COVID-19 has high infectivity, strong pathogenicity, and a long incubation period. During the treatment of COVID-19 patients, the disease changed rapidly. Multiple re-examinations were required in a short period, which generated much medical data. These problems make the accurate and rapid diagnosis of COVID-19 extremely difficult for clinicians. AI can quickly analyze large quantities of test data to find suspected COVID-19 patients. It also evaluates the degree of lung damage, compares the patient's condition before and after the illness in detail, and quickly assesses the

disease progress. These works accurately assist the doctor in diagnosis (30,31). This section focuses on AI applications in diagnosis areas, such as laboratory-based diagnosis, medical image diagnosis, respiratory pattern, and symptoms diagnosis (Table 2).

Laboratory-based diagnosis

The standard diagnostic approach for COVID-19 is the real-time reverse-transcriptase polymerase chain reaction (rRT-PCR) technique with DNA sequencing and identification. Gomes *et al.* (32) proposed a pseudo-convolutional machine learning method to improve the process of DNA identification by dividing the DNA sequence into more minor sequences with overlap. Then it optimized the COVID-19 molecular diagnosis to identify SARS-Cov-2 DNA sequences faster with higher specificity and sensitivity by different models, such as random forests (RF), naive Bayes classifier (NBC), instance-based learner (IBL), multilayer perceptron (MLP), support vector machine (SVM). For example, experiments with all 24 virus families and SARS-Cov-2 resulted in 0.822222 ± 0.05613 for sensitivity and 0.99974 ± 0.00001 for specificity using RF with 100 trees and 30% overlap. Villarreal-González *et al.* (40) detected typical profiles in PCR curves caused by contamination or artifacts. So, they compared the accuracy and log loss parameters from different ML methods, such as K-neighbor classifier, SVM, decision tree classifier, quadratic discriminant analysis (QDA), linear discriminant analysis (LDA), and RF. LDA presented the highest accuracy value of 97.6 and a low log loss of 0.1. It was the best method to optimize PCR tests' results. Data of antibodies, antigens, and routine blood exams were also used to identify the COVID-19 cases. Cady *et al.* (33) used a multiplexed grating-coupled fluorescent plasmonics (GC-FP) biosensor platform to rapidly and accurately measure COVID-19 antibodies in human blood serum. A machine learning approach based on SVM was used to analyze GC-FP detection data. The measure antibodies against COVID-19 in human blood serum and dried blood spot samples were 100% and 86.7% for sensitivity. Kukar *et al.* (34) constructed a CRISP-deep neural network (DNN) based on routine blood tests to discriminate COVID-19 patients from patients with other infectious diseases. According to the feature importance scoring of the eXtreme Gradient Boosting (XGBoost) model, it used five usual blood parameters, which were mean corpuscular hemoglobin concentration (MCHC), eosinophil count,

albumin, international normalized ratio (INR), and prothrombin activity percentage. The sensitivity, specificity, and area under the curve (AUC) were 81.9%, 97.9%, and 0.97, respectively. The results were an impressively low proportion of false positives and a moderately low proportion of false negatives. The model was helpful in the early symptomatic phase when COVID-19 was easier to be missed by the RT-PCR test.

Medical images diagnosis

The advantages of imaging examination are intuitive, fast, and accurate, making it an essential basis for clinical diagnosis. The most common imaging examinations for the COVID-2019 diagnosis and treatment are CT examination and chest X-ray (41-44). Due to high resolution, low missed diagnosis rate, and better display of ground-glass opacity (GGO), CT image has become one of the gold diagnosis standards in the COVID-2019 Pandemic. Multiple CT diagnosis results are classified as binary classification problems, such as COVID-19 negative or COVID-19 positive. Using the CT images of 723 COVID-19 positive cases and 1,145 negative cases, Wang *et al.* (35) proposed a CNN-based model to determine the COVID-19 cases. The experimental results showed that the proposed model's sensitivity was 97.4%, the specificity was 92.2%, and the AUC was 0.991. It would save about 30–40% of the check time for physicians and promote the performance of COVID-19 detection. When AI technology processes CT images of COVID-19, segmentation is an indispensable step in image processing and analysis. By segmenting and extracting regions of interest (ROI), such as lungs, lung lobes, bronchopulmonary segments, and infected or diseased areas, AI can further evaluate and quantify, helping solve the shortage and distribution of human medical resources. Wang *et al.* (35) proposed three processes for screening COVID-19 in CT images. First, it detected the entire lung area through an effective segmentation network based on UNet++; then, it located the lesion or nodule area; finally, it isolated and found the lung lesion area. Common pneumonia radiological images, especially viral pneumonia, are similar to COVID-19. Other machine learning methods attribute diagnosis to three classification tasks: health, COVID-19 patients, and different types of pneumonia patients. Xu *et al.* (36) used the chest CT images of 110 COVID-19 patients, 224 influenza A patients, and 175 healthy people to classify and utilized a V-Net-based deep learning model to segment candidate infection areas.

Table 2 Applications of AI in COVID-19 diagnosis

First author [year] (reference)	Country (region)	Modality	Model	Data source	Sample size	Application area	Result
Gomes <i>et al.</i> [2020] (32)	Brazil	DNA sequences	RF/NBC/IBL/MLP/SVM	NIAID Virus Pathogen Database/Analysis Resource (VIPR)/Genome Reference Consortium	Twenty-five different viruses from NIAID Virus Pathogen Database/Analysis Resource (VIPR): 347,363; viruses from Genome Reference Consortium: 103,959	Laboratory-based diagnosis: RT-PCR method	RF (results from dataset with 30% overlap) sensitivity: 0.822222±0.05613; specificity: 0.99974±0.00001; AUC: 0.99884±0.0025; MLP (results from dataset with 30% overlap) sensitivity: 0.97386±0.03052; specificity: 0.96151±0.00246; AUC: 0.97353±0.01863; MLP (results from dataset with 50% overlap) sensitivity: 0.98824±0.01198; specificity: 0.99860±0.00020; AUC: 0.99947±0.00056
Cady <i>et al.</i> [2021] (33)	USA	Blood sample	SVM	Dataset: COVID-19 negative samples plus COVID-19 positive samples	Negative samples: obtained from the Lyme Disease Biobank prior to the COVID-19 pandemic; positive samples: obtained from donors within New York state or from the Wadsworth Center, New York State Department of Health	Laboratory-based diagnosis: antibody response	Selectivity of human blood serum: 100%, sensitivity of dried blood spot samples: 86.7%
Kukar <i>et al.</i> [2021] (34)	Switzerland	Blood sample	CRISP-DM based machine learning/DNN/RF/XGBoost	Dataset: COVID-19 negative cases plus COVID-19 positive cases	Healthy cases: 5,108 from March, 2012 to April, 2019; COVID-19 cases: 160 from March/April 2020; other cases (different bacterial and viral infections): 225 from March, 2012 to April, 2019	Laboratory-based diagnosis: routine blood tests	Sensitivity: 81.9%, specificity: 97.9%, AUC: 0.97
Wang <i>et al.</i> [2021] (35)	China	CT image	3D U-Net++/residual network (ResNet-50)	Training dataset: 1,136 (723 were positive); testing dataset: 282 cases (154 were positive)	Healthy cases: 70; COVID-19 cases: 723; other cases (inflammation or tumors): 343	Medical images diagnosis: CT image	Sensitivity: 0.974, specificity: 0.922, AUC: 0.991

Table 2 (continued)

Table 2 (continued)

First author [year] (reference)	Country (region)	Modality	Model	Data source	Sample size	Application area	Result
Xu et al. [2020] (36)	China	CT image	3D CNN segmentation model/residual network (ResNet-18)	Training dataset: 528 CT samples (COVID-19: 189; IAVP: 194; healthy: 145); testing dataset: 90 CT samples (COVID-19: 30; IAVP: 30; healthy: 30)	Healthy cases/CT samples: 175/175; COVID-19 cases/CT samples: 110/219; other cases IAVP/CT samples: 224/224	Medical images diagnosis: CT image	Accuracy: 86.7%
Narin et al. [2021] (37)	Globe	Chest X-ray radiographs	CNN based models (ResNet50, ResNet101, ResNet152, InceptionV3 and Inception-ResNetV2)	Dataset-1: 341 (COVID-19)/2,800 (healthy); dataset-2: 341 (COVID-19)/1,493 (viral); dataset-3: 341 (COVID-19)/2,772 (bacterial)	Healthy X-ray samples: 2,800; COVID-19 X-ray samples: 341; other viral pneumonia X-ray samples: 1,493; other bacterial pneumonia X-ray samples: 2,772	Medical images diagnosis: chest X-ray radiographs	Accuracy for dataset-1: 96.1%, accuracy for dataset-2: 99.5%, accuracy for dataset-3: 99.7%
Zhang et al. [2020] (38)	Globe	Chest X-ray radiographs	CAAD model	Dataset-1 cases/X-ray samples: 33/50 (COVID-19); 492/714 (other pneumonia); dataset-2 cases/X-ray samples: 37/50 (COVID-19); 516/717 (other pneumonia)	COVID-19 cases/X-ray samples: 70/100; other pneumonia cases/X-ray samples: 1,008/1,431	Medical images diagnosis: chest X-ray radiographs	Sensitivity: 96.0%, specificity: 70.7% AUC: 0.952
Arpaci et al. [2021] (39)	China	Clinical features	BayesNet/Logistic/IBk/CR/PART/J48	114 subjects from the Taizhou hospital of Zhejiang Province in China from January 17, 2020 to February 1, 2020	COVID-19 cases: 32; non COVID-19 cases: 82	Respiratory pattern and symptoms diagnosis	Accuracy of BayesNet: 71.93%; accuracy of Logistic: 80.7%; accuracy of IBk: 72.81%; accuracy of CR: 84.21%; accuracy of PART: 76.32%; accuracy of J48: 73.68%

AI, artificial intelligence; COVID-19, coronavirus disease 2019; RF, random forests; IBL, naive Bayes classifier; IBL, instance-based learner; MLP, multilayer perceptron; SVM, support vector machine; DNN, deep neural networks; XGBoost, extreme gradient boosting machine; CAAD, confidence-aware anomaly detection model; BayesNet, Bayes classifier; Logistic, logistic-regression; IBk, lazy-classifier; PART, rule-learner; J48, decision-tree; IAVP, influenza-A viral pneumonia; RT-PCR, reverse-transcriptase polymerase chain reaction; AUC, area under the curve.

The relative infection distance characteristics of patches and edges were sent to the ResNet-18 network together, and the output was one of these three groups. The overall accuracy of the model was 86.7%.

X-ray has the advantages of continuous monitoring, low radiation dose, and easy operation. Narin *et al.* (37) proposed five different CNN-based models to implement three different binary classifications with COVID-19 patients, healthy cases, and other pneumonia cases from X-ray images. They used data from the various online dataset, such as the open-source GitHub dataset and the Kaggle pneumonia dataset. The evaluation results showed that ResNet50 had an accuracy of 98.0%, achieving the highest classification performance. Zhang *et al.* (38) proposed a ResNet-based model to detect COVID-19 in X-ray images. The model had two functions. Not only it classified cases in COVID-19 or non-COVID-19, but also it was used for anomaly detection. The anomaly detection task would give an anomaly score to optimize the COVID-19 score for classification. The two data sets include X-ray images from 70 COVID-19 patients and 1,008 non-COVID-19 patients. The sensitivity and specificity were 96.0% and 70.7%, respectively, and the AUC was 0.952.

Respiratory pattern and symptoms diagnosis

Another diagnostic sign of COVID-19 patients was their different respiratory pattern from common cold and influenza (45). Abnormally rapid breathing was an essential indication of COVID-19 infection. Therefore, Wang *et al.* (46) developed a COVID-19 diagnostic method based on respiratory features. They applied a GRU neural network with bidirectional and attentional mechanisms (BI-AT-GRU) to classify 6 clinically significant respiratory patterns (Eupnea, Tachypnea, Bradypnea, Biots, Cheyne-Stokes, and Central- Apnea). The results showed that the proposed model could classify six different respiratory patterns with the accuracy, precision, recall, and F1 of 94.5%, 94.4%, 95.1%, and 94.8%, respectively. Arpacı *et al.* (39) analyzed 114 cases from the Taizhou hospital of Zhejiang Province in China. They developed six predictive models for COVID-19 diagnosis using six different classifiers based on 14 clinical features, including Bayes classifier (BayesNet), logistic-regression (Logistic), lazy-classifier (IBk), meta-classifier [classification via regression (CR), rule-learner (PART), and decision-tree (J48)]. As a result, the CR was the most accurate classifier for predicting the positive or negative COVID-19 cases with an accuracy of 84.21%. Mouawad

et al. (47) employed symbolic recurrence quantification measures with MFCC features to detect healthy people's sustained vowels or sick individuals' cough sounds. The proposed model achieved a mean classification performance of 97% and 99%, and a mean F1-score of 91% and 89% after optimization, for coughs and sustained vowels, respectively. These studies could help medical workers to diagnose COVID-19 patients more effectively and reduce the current medical system's detection workload.

Applications of AI in the COVID-19 progression

The AI screening research can detect early COVID-19 cases and improve the doctors' diagnosis. Similarly, AI applications in the disease progression are also critical, helping medical staff find and treat high-risk patients early, estimate ICU events, formulate treatment plans, allocate medical resources, and reduce mortality (48-50) (Table 3). Li *et al.* (51) investigated the AI-assisted quantification on COVID-19 patients' initial chest CT to predict their disease progression and clinical outcome. In this study, they calculated the CT severity score (CT-SS) according to the extent of lesion involvement. AI-based quantification of GGO and consolidation volume were performed. Among imaging parameters, consolidation volume had the largest AUC in discriminating non-severe from progress-to-severe group (AUC =0.796, P<0.001) and patients with or without critical events (AUC =0.754, P<0.001). The results indicated that consolidation volume and age were the two strongest predictors for disease progression.

Moreover, the larger consolidation volume was associated with unfavorable clinical outcomes. Yang *et al.* (52) applied chest CT-SS as an imaging tool for assessing COVID-19 progression. The optimal CT-SS threshold for identifying severe COVID-19 was 19.5 (AUC =0.892), with 83.3% sensitivity and 94% specificity. CT-SS could evaluate the severity of pulmonary involvement quickly and objectively in COVID-19 patients. Yan *et al.* (53) developed a predictive model based on the XGBoost model. They identified three vital clinical features: lactic dehydrogenase, lymphocyte count, and high sensitivity C-reactive protein, from more than 300 features. The model could predict the survival rate of COVID-19 patients, with an accuracy rate exceeding 90%.

Applications of AI in the COVID-19 treatment

So far, there is no cure for COVID-19 disease. The

Table 3 Applications of AI in COVID-19 progression

First author [year] (reference)	Country (region)	Modality	Model	Data source	Sample size	Result
Li et al. [2020] (51)	China	CT image	U-Net	COVID-19 patients in Shanghai Jiao Tong University Affiliated Sixth People's Hospital from February 10, 2020 to April 9, 2020	COVID-19 cases classified as non-severe group on admission: 123	(CT-SS) AUC 0.66; accuracy 62.6%; sensitivity 58.97%; specificity: 64.29% (GGO volume cm ³) AUC 0.639; accuracy 43.9%; sensitivity 79.49%; specificity: 45.24% (GGO volume percentage): AUC 0.694; accuracy 62.6%; sensitivity 64.1%; specificity: 69.05%; (consolidation volume cm ³): AUC 0.796; accuracy 78.05%; sensitivity 71.79%; specificity: 80.95%; (consolidation volume percentage): AUC 0.79; accuracy 78.86%; sensitivity 79.49%; specificity: 78.57%
Yang et al. [2020] (52)	China	CT image	CT-SS	COVID-19 patients in Chongqing Three Gorges Central Hospital from January 21, 2020 to February 5, 2020	COVID-19 cases: 102	AUC: 0.892; sensitivity: 83.3%; specificity: 94%

AI, artificial intelligence; COVID-19, coronavirus disease 2019; CT-SS, CT severity score; AUC, area under the curve; GGO, ground-glass opacity.

development of vaccines and drugs, coupled with corresponding valuable clinical treatment plans, is the ultimate means to solve the COVID-19 Pandemic. AI can apply to the four stages of drug development: drug discovery, preclinical research, clinical research, and marketing approval. It can also reduce the repetitive work in virus detection to accelerate the development of vaccines by effectively screening compounds, biomarkers and predicting drugs' physical and chemical properties. This section focuses on AI applications in the COVID-19 treatments, such as drug designing, drug repurposing, herbal drugs, and vaccine development (Table 4).

Drug designing

The 3D structure of the protein encoded by the SARS-CoV-2 gene is used as the main or potential drug treatment target. Generally, the protein structure can be determined by experimental methods such as X-ray crystal diffraction spectroscopy. But these methods are expensive and time-consuming. AI screening methods can predict the structure of these related proteins, identify drugs showing high affinity for different target proteins, and propose new chemical compounds as potential treatments (60,61). Pfab *et al.* (54) applied a fully automated deep learning-based method called DeepTracer, whose central piece was the CNN. By predicting four vital pieces of information (the locations of amino acids, the location of the backbone, secondary structure positions, and amino acid types), it could determine the *de novo* multichain protein complex structure from high-resolution cryoelectron microscopy maps fastly. The average percentage of matched model residues was 84% for DeepTracer, and it also achieved a sequence matching percentage of 63.08%. This Information about the macromolecular structure of protein complexes could assist the development of vaccines and drugs. Magar *et al.* (55) proposed different machine-learning models such as XGBoost, RF, MLP, SVM, and LR for high throughput screening of synthetic antibodies to discover the possible inhibitory antibodies for COVID-19. They trained the ML models with 14 different virus types and achieved over 90% fivefold test accuracy. The models predicted the antibodies neutralization, found 18 antibodies highly efficient in neutralizing SARS-CoV-2. Moreover, it checked the stability of predicted antibodies and found nine stable antibodies which could neutralize SARS-CoV-2.

Table 4 Applications of AI in COVID-19 treatment

First author [year] (reference)	Country (region)	Modality	Model	Data source	Application area	Result
Prab <i>et al.</i> [2021] (54)	USA	Multichain protein complex structure	CNN/U-Net	EM Data Resource/Protein Data Bank	Drug designing	The average percentage of matched model residues: 84%; the sequence matching percentage: 63.08%
Megar <i>et al.</i> [2021] (55)	USA	Amino acid sequences	XGBoost/RF/MLP/SVM/LR	Virus Net dataset	Drug designing	Accuracy: XGBoost (90.57%)/RF (89.18%)/LR (81.17%)/MLP (78.23%)/SVM (75.49%)
Beck <i>et al.</i> [2020] (56)	Korea	Amino acid sequences	MT-DTI/NLP	NCBI database/DTC database/Binding DB database	Drug repurposing	Predict drug-target interactions accurately and find valuable drugs for COVID-19
Zeng <i>et al.</i> [2020] (57)	China	PubMed publications	KG-DML	Global Network of Biomedical Relationships/Drug Bank/transcriptome datasets	Drug repurposing	AUC: 0.85
Erlina <i>et al.</i> [2020] (58)	Indonesia	Multichain protein complex structure	SVM/RF/MLP	Super Target web resources/Herbal database	Herbal drug	Accuracy: SVM (99.91%)/RF (98.67%)/MLP (98.32%); AUC: SVM (0.99)/RF (0.98)/MLP (0.98); precision: SVM (99.84%)/RF (97.25%)/MLP (96.63%)
Ong <i>et al.</i> [2020] (59)	USA	PubMed publications/multichain protein complex structure/DNA sequences	RF/SVM	PubMed database/ClinicalTrials.gov database	Vaccines development	Six proteins, including the S protein and five non-structural proteins (Nsp3, 3CL-pro, and Nsp8-10), were predicted to be adhesins, which are crucial to the viral adhering and host invasion

AI, artificial intelligence; COVID-19, coronavirus disease 2019; CNN, convolutional neural networks; XGBoost, extreme gradient boosting machine; RF, random forests; MLP, multilayer perceptron; SVM, support vector machine; LR, logistic regression; MT-DTI, molecule transformer-drug target interaction; NLP, natural language processing; KG-DML, knowledge-graph-based deep-learning model; NCBI, National Center for Biotechnology Information; DTC, Drug Target Common.

Drug repurposing

Drug repurposing is a technology that uses existing drugs to treat emerging and challenging diseases to reduce development time and overall costs. Among the several treatment approaches, drug repurposing has an edge in causing few side effects, making it a promising candidate for developing new therapeutic strategies. Therefore, it is urgent to use AI technology to discover drugs with inhibitory effects on the SARS-CoV-2 from existing drugs (62-64). There was much complex information between molecules, such as hydrophobic interactions, ionic interactions, bonding between hydrogen molecules, or van der Waals forces. Beck *et al.* (56) used a pre-trained deep learning model to predict drug-target interactions accurately. The model's core is natural language processing (NLP) based on the bidirectional encoder representations transformers (BERT) framework, which had good performance and reliability in various drug-targeted interaction data sets. They analyzed the COVID-19 3C-like proteinase and 3,410 existing drugs available in the market. A popular antiretroviral drug used to treat HIV called Antazanavir (Kd of 94.94 nM) was the best drug for COVID-19 medication, followed by efavirenz (Kd value of 199.17 nM), ritonavir (Kd value of 204.05 nM), and dolutegravir (Kd value of 336.91 nM). Zeng *et al.* (57) proposed an integrative, network-based deep-learning methodology to analyze many medical literature and related data, quickly discovered the connections between drugs and diseases, diseases and genes, and established the comprehensive knowledge graph. The graph included 15 million edges across 39 relationships connecting drugs, diseases, pathways, genes, and expressions from 24 million publications. Using the ongoing COVID-19 trial data as a validation set, this model had a larger AUC (0.85) for identifying repurposable drugs. Forty-one repurposable drugs (including dexamethasone, indomethacin, niclosamide, and toremifene) with inhibitory effects on the SARS-CoV-2 were identified and validated.

Herbal drug

Due to the unclear basis and mechanism of active substances, the discovery, use, and promotion of herbal medicines have been greatly restricted. Some researchers found particular effectiveness in treating COVID-19 cases with the herbal drug (65,66). Liu *et al.* (67) constructed a compounds-herb-target organ-function network by the

computer-aided drug design (CADD) based molecular docking technology. They identified potential SARS-CoV-2 3CL protease inhibitors from the traditional Chinese medicine systems pharmacology database and analysis platform (TCMSP). This work revealed that, to a certain extent, Glycyrrhizae Radix et Rhizoma, Rhododendron dauricum, and Plantaginis Herba, etc. had capabilities to relieve cough or asthma and dispel lung-draining. Erlina *et al.* (58) adopted SVM, MLP, and RF models to study the Indonesian herbal compound and its effectiveness. The accuracy and f-measure of the model of each method were high, around 98%, respectively. The authors adopted the structure-based method for pharmacophore modeling, which combined the 3D structure of COVID-19 main protease. The result showed that six herbal compounds, i.e., Hesperidin, Kaempferol-3,4'-di-O-methyl ether (Ermanin); Myricetin-3-glucoside, Peonidine 3-(4'-arabinosylglucoside); Quercetin 3-(2G-rhamnosylrutinoside); and Rhamnetin 3-mannosyl-(1,2)-alloside, could serve as effective COVID-19 drugs.

Vaccines development

To resist the high viral infection rate, it is essential to determine the best target for developing a vaccine. The host immune system fights virus-infected cells by B cells producing antibodies or the direct attack of T cells. HLA genes encode MCH-I and MCH-II proteins, which present epitopes as antigenic determinants. These proteins assist the ability of B cells and T cell antibodies to bind and attack invaders. Machine learning methods, including RF, SVM, and RFE, have become essential tools for identifying antigens from protein sequences. However, due to their low sensitivity in predicting local clustering interactions in some cases, deep CNN (DCNN) have always been a more practical choice for the combined prediction of MHC and peptides (68-70). Ong *et al.* (59) used the newly developed machine learning model and Reverse Vaccinology (RV) tools to predict COVID-19 vaccine candidates. Besides the commonly used structural protein (Sp), they prioritized non-Sps as vaccine candidates for SARS-CoV-2. As the most significant non-Sp in the coronavirus family, non-Sp 3 (Nsp3) was considered the most promising vaccine development target after Spike. The "Sp/Nsp cocktail vaccine" containing an Sp and an Nsp would stimulate effective complementary immune responses. Malone *et al.* (71) also studied the entire SARS-CoV-2 proteome except for Spike, identified a subset of epitope hotspots

that could be harnessed in a vaccine formulation, and used NEC immunoassay analyzer, IEDB, and BepiPred tools, demonstrating accuracy of 87%, to establish epitope maps of different HLA alleles. It provided a SARS-CoV-2 comprehensive vaccine design blueprint.

Applications of AI in the COVID-19 psychological effects

COVID-19 Pandemic had unprecedented and far-reaching impacts on mental health. Severe SARS-CoV-2 infection may produce mental symptoms for the following reasons: direct effects caused by viral infection (such as hypoxemia), immune response, and medical intervention. Other explanations involve broader social impacts, including psychological impacts due to social isolation, unemployment, the fear of getting infected, inadequate psychological support, racial discrimination, or the psychological burden caused by fear of infecting others (72,73) (Table 5). Choi *et al.* (74) used the ANN model to research the mental health effects of racial discrimination targeting the Asian population in the U.S. during the COVID-19 Pandemic. They unveiled the most critical factors affecting Korean immigrants' psychological distress, such as individuals' level of resilience, the experience of everyday discrimination, and intensified racial discrimination toward Asian populations. The model's AUC was 0.806. Wang *et al.* (75) implemented XGBoost models to investigate the prevalence and severity of anxiety among Chinese non-graduating college students. It also compared the difference between the anxiety status 1 month before and after starting the new semester of online learning during COVID-19. The accuracy rate was approximately 80%. It could help related departments implement adequate measures, such as timely psychological intervention before the anxiety levels increase. They used Integrating Bayesian networks with classical machine learning to identify factors that significantly impacted mental health during COVID Pandemic (76). The mode analyzed 17,764 adults in the USA at different age groups, genders, and socioeconomic statuses. They observed the accuracy in model predictability decayed from 0.80 to 0.64 when they moved from high risk of depression to low risk of depression. Overall, females were more stressed than males, and people of age-group 18–29 were more vulnerable to anxiety than other age groups. Ćosić *et al.*, conducted a comprehensive approach to address the lack of access to psychiatric services. The process included AI, telepsychiatry, and a range of novel

technologies, like internet-based computer-aided mental health tools (78).

AI in medical data during COVID-19 pandemic

In the era of big data, as the main body of responding to public emergencies, governments of various countries used big data to control the COVID-19 Pandemic in real-time. AI needs to store a large amount of medical data to analyze, most of which are patients' private medical information, such as the medical history of present disease, genetic history, physical defects, treatment conditions, etc. With the increasing awareness of people's privacy protection and data security, striving to strike a balance between the three social goals of digital technology advancement, patient privacy protection, and public health maintenance is an eternal topic that needed attention at this time (79,80) (Table 5). Social media like Facebook, Twitter, YouTube, Instagram, Snapchat, and WhatsApp were the primary source for spreading information and news in the COVID-19 period. Facebook also shared anonymized population flow data and population density summary maps with other scientific research institutions to build the COVID-19 infection chain model (81). Kang *et al.* (77) developed a web-based data AI platform to share massive pathological image data. This platform included 3,100 images for AI researchers to quickly load images into their learning models. Several precautions were followed to ensure that any information potentially identifying subjects, including the patients, was not contained in the collected dataset. After submitting the consent forms for the use of data ["data use agreement" (DUA)], researchers receiving approval could download the data. However, to prevent reckless data leakage and usefulness, users were managed at four levels based on their right to use the data. Kumar *et al.* (82) presented a deep learning model to predict lung cancer in the healthcare system by filling the defined gap. They proposed a method to secure medical data by only sharing the weights of the trained deep learning model via smart contract.

Discussion

As an empirical science, medicine itself has many uncertainties. AI in the medical field still has the following problems worthy of our consideration.

- (I) From a technical perspective, AI development requires many shared data platforms composed of a large amount of data. Different data needs

Table 5 Applications of AI in COVID-19 psychological effects and data security

First author [year] (reference)	Country (region)	Modality	Model	Data source	Application area	Result
Choi <i>et al.</i> [2020] (74)	USA	Sociodemographic questionnaire	ANN	Korean immigrants above the age of 18 residing in the U.S. were invited to respond to a survey by e-mails and posting on Korean immigrants' online communities from 24 May 2020 to 14 June 2020	Psychological effects	AUC: 0.806
Wang <i>et al.</i> [2020] (75)	China	Sociodemographic questionnaire	XGBoost	3,800 non-graduating college students from a top multidisciplinary and research-oriented university directly under the jurisdiction of the Ministry of Education in North China were invited to attend the studies during February 15 to March 17, 2020	Psychological effects	Accuracy of Model 1: 79.26%; accuracy of Model 2: 84.38%
Jha <i>et al.</i> [2020] (76)	USA	Sociodemographic questionnaire	PGM	17,764 adults in the USA at different age groups, genders, and socioeconomic statuses	Psychological effects	Accuracy of high risk of depression group: 0.80; accuracy of low risk of depression group: 0.64
Kang <i>et al.</i> [2021] (77)	Korea	Pathological image data	PAIP	3,100 images acquired by the Department of Pathology at Seoul National University Hospital, Seoul National University Bundang Hospital, and SMG-SNU Boramae Medical Center	Data security	Accuracy of liver cancer: 83%; accuracy of prostate cancer: 86%; accuracy of kidney cancer: 80%

AI, artificial intelligence; COVID-19, coronavirus disease 2019; ANN, artificial neural network; XGBoost, extreme gradient boosting machine; PGM, Bayesian probabilistic graphical model; PAIP, pathology artificial intelligence platform.

to be unified in format, type, and label code. The continuous growth of data requires incremental learning methods. How to reduce the noise of the data set is also essential. We need to promote and use the related intelligent devices. AI will be popularized when the technological problems were solved (83,84).

- (II) From the perspective of clinical applications, although black box models such as neural networks, DNNs, deep learning, or gradient enhancement models have high accuracy, their internal working mechanisms are complicated for medical staff to understand. It is difficult for medical users to find product errors after long-term use. The front-line clinical team has the most profound and direct feelings about the prevention and control of diseases. They can dig out professional needs from medical practice, carry out clinical trials, collect adequate evidence, and evaluate AI technology. They and AI researchers can integrate AI technology with current clinical processes and information systems into applications. In the future, AI personnel and medical workers will further cooperate closely (85,86).
- (III) From the perspective of data safety, AI learning requires more high-quality data for training. Due to the complex characteristics and the unclear property relationship of medical data, fewer data meet the training conditions. The medical data involve patient's privacy, so it is easy to cause leakage of patient data. How to protect patient's privacy will become a focal part of AI research (77).

In the process of AI application, what changes will medical care have in the future? China has a large population, vast land, and abundant resources. The medical market is enormous and rich. How to seize the opportunity to promote AI medical technology to a world-class level is a big challenge. More and more AI medical processes begin to pay attention to integrity and integration, from a single disease to multiple diseases, from single information to numerous information. AI technology will apply in the whole life cycle of health management, including pre-diagnosis, during-diagnosis, and post-diagnosis. It integrates the prevention, diagnosis, treatment, rehabilitation, follow-up, and other links. We need to do more to prevent disease before the disease epidemic. Medical staff, including clinicians, disease control personnel, clinical auxiliary departments, and administrative management personnel, will

become the leading force of AI medical care in the future.

Acknowledgments

Funding: This study was supported by the Special Innovation Projects of Ordinary Universities in Guangdong Province, project name: Research on Visual Recognition and Sorting Robots Based on Deep Learning (2018KTSCX260); Project of Educational Science Planning of Guangdong Province in China: Research on the Construction of Intelligent Robot College in Guangdong, Hong Kong, and Macao (2020WQYB014).

Footnote

Provenance and Peer Review: This article was commissioned by the Guest Editors (Jianxing He and Hengrui Liang) for the series "Artificial Intelligence in Thoracic Disease: from Bench to Bed" published in *Journal of Thoracic Disease*. The article has undergone external peer review.

Reporting Checklist: The authors have completed the PRISMA reporting checklist. Available at: <https://dx.doi.org/10.21037/jtd-21-747>

Peer Review File: Available at <https://dx.doi.org/10.21037/jtd-21-747>

Conflicts of Interest: The authors have completed the ICMJE uniform disclosure form (available at <https://dx.doi.org/10.21037/jtd-21-747>). The series "Artificial Intelligence in Thoracic Disease: from Bench to Bed" was commissioned by the editorial office without any funding or sponsorship. WL serves as an unpaid editorial board member of *Journal of Thoracic Disease*. The authors have no other conflicts of interest to declare.

Ethical Statement: The authors are accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved.

Open Access Statement: This is an Open Access article distributed in accordance with the Creative Commons Attribution-NonCommercial-NoDerivs 4.0 International License (CC BY-NC-ND 4.0), which permits the non-commercial replication and distribution of the article with the strict proviso that no changes or edits are made and the

original work is properly cited (including links to both the formal publication through the relevant DOI and the license). See: <https://creativecommons.org/licenses/by-nc-nd/4.0/>.

References

- Budd J, Miller BS, Manning EM, et al. Digital technologies in the public-health response to COVID-19. *Nat Med* 2020;26:1183-92.
- Guan WJ, Ni ZY, Hu Y, et al. Clinical Characteristics of Coronavirus Disease 2019 in China. *N Engl J Med* 2020;382:1708-20.
- Chan JF, Yuan S, Kok KH, et al. A familial cluster of pneumonia associated with the 2019 novel coronavirus indicating person-to-person transmission: a study of a family cluster. *Lancet* 2020;395:514-23.
- Harrison AG, Lin T, Wang P. Mechanisms of SARS-CoV-2 Transmission and Pathogenesis. *Trends Immunol* 2020;41:1100-15.
- Wang MY, Zhao R, Gao LJ, et al. SARS-CoV-2: Structure, Biology, and Structure-Based Therapeutics Development. *Front Cell Infect Microbiol* 2020;10:587269.
- Kinoshita R, Anzai A, Jung SM, et al. Containment, Contact Tracing and Asymptomatic Transmission of Novel Coronavirus Disease (COVID-19): A Modelling Study. *J Clin Med* 2020;9:3125.
- Bragazzi NL, Dai H, Damiani G, et al. How Big Data and Artificial Intelligence Can Help Better Manage the COVID-19 Pandemic. *Int J Environ Res Public Health* 2020;17:3176.
- Naseem M, Akhund R, Arshad H, et al. Exploring the Potential of Artificial Intelligence and Machine Learning to Combat COVID-19 and Existing Opportunities for LMIC: A Scoping Review. *J Prim Care Community Health* 2020;11:2150132720963634.
- Zanin M, Xiao C, Liang T, et al. The public health response to the COVID-19 outbreak in mainland China: a narrative review. *J Thorac Dis* 2020;12:4434-49.
- Aslan IH, Demir M, Wise MM, et al. Modeling COVID-19: forecasting and analyzing the dynamics of the outbreak in Hubei and Turkey. *medRxiv* 2020. doi: 10.1101/2020.04.11.20061952.
- Yang Z, Zeng Z, Wang K, et al. Modified SEIR and AI prediction of the epidemics trend of COVID-19 in China under public health interventions. *J Thorac Dis* 2020;12:165-74.
- Chimmula VKR, Zhang L. Time series forecasting of COVID-19 transmission in Canada using LSTM networks. *Chaos Solitons Fractals* 2020;135:109864.
- Kolozsvari LR, Berczes T, Hajdu A, et al. Predicting the epidemic curve of the coronavirus (SARS-CoV-2) disease (COVID-19) using artificial intelligence. *medRxiv* 2021. doi: 10.1101/2020.04.17.20069666.
- Hu Z, Ge Q, Li S, et al. Artificial intelligence forecasting of covid-19 in China. *arXiv* 2020:2002.07112.
- Fong SJ, Li G, Dey N, et al. Finding an accurate early forecasting model from small dataset: a case of 2019-ncov novel coronavirus outbreak. *arXiv* 2020:2003.10776.
- Mao Z, Yao H, Zou Q, et al. Digital Contact Tracing Based on a Graph Database Algorithm for Emergency Management During the COVID-19 Epidemic: Case Study. *JMIR Mhealth Uhealth* 2021;9:e26836.
- Pramanik M, Chowdhury K, Rana MJ, et al. Climatic influence on the magnitude of COVID-19 outbreak: a stochastic model-based global analysis. *Int J Environ Health Res* 2020. [Epub ahead of print]. doi: 10.1080/09603123.2020.1831446.
- Pasayat AK, Pati SN, Maharana A. Predicting the COVID-19 positive cases in India with concern to Lockdown by using Mathematical and Machine Learning based Models. *medRxiv* 2020. doi: 10.1101/2020.05.16.20104133.
- Kermack WO, McKendrick AG. A contribution to the mathematical theory of epidemics. *Proc R Soc Lond* 1927;115:700-21.
- Dandekar R, Barbastathis G. Neural Network aided quarantine control model estimation of COVID spread in Wuhan, China. *arXiv* 2020:2003.09403.
- Al-Raei M. The forecasting of COVID-19 with mortality using SIRD epidemic model for the United States, Russia, China, and the Syrian Arab Republic. *Aip Advances* 2020;10:065325.
- Rumelhart DE, Hinton GE, Williams RJ. Learning representations by back-propagating errors. *Nature* 1986;323:533-6.
- Adly AS, Adly AS, Adly MS. Approaches Based on Artificial Intelligence and the Internet of Intelligent Things to Prevent the Spread of COVID-19: Scoping Review. *J Med Internet Res* 2020;22:e19104.
- Srinivasa Rao ASR, Vazquez JA. Identification of COVID-19 can be quicker through artificial intelligence framework using a mobile phone-based survey when cities and towns are under quarantine. *Infect Control Hosp Epidemiol* 2020;41:826-30.
- Wang CJ, Ng CY, Brook RH. Response to COVID-19 in Taiwan: Big Data Analytics, New Technology, and

- Proactive Testing. *JAMA* 2020;323:1341-2.
26. Travaglio M, Yu Y, Popovic R, et al. Links between air pollution and COVID-19 in England. *Environ Pollut* 2021;268:115859.
 27. Mishra AK, Wargocki P. Attempting to understand the spread of SARS-CoV-2: Associations between the number of cases in regions and selected economic and geographic parameters. *arXiv* 2020:2004.00110.
 28. Zhang X, Ji Z, Zheng Y, et al. Evaluating the effect of city lock-down on controlling COVID-19 propagation through deep learning and network science models. *Cities* 2020;107:102869.
 29. Vaid S, McAdie A, Kremer R, et al. Risk of a second wave of Covid-19 infections: using artificial intelligence to investigate stringency of physical distancing policies in North America. *Int Orthop* 2020;44:1581-9.
 30. Li WT, Ma J, Shende N, et al. Using machine learning of clinical data to diagnose COVID-19: a systematic review and meta-analysis. *BMC Med Inform Decis Mak* 2020;20:247.
 31. Brinati D, Campagner A, Ferrari D, et al. Detection of COVID-19 Infection from Routine Blood Exams with Machine Learning: A Feasibility Study. *J Med Syst* 2020;44:135.
 32. Gomes JC, Masood AI, Silva LHS, et al. Optimizing the molecular diagnosis of Covid-19 by combining RT-PCR and a pseudo-convolutional machine learning approach to characterize virus DNA sequences. *bioRxiv* 2020. doi: 10.1101/2020.06.02.129775.
 33. Cady NC, Tokranova N, Minor A, et al. Multiplexed detection and quantification of human antibody response to COVID-19 infection using a plasmon enhanced biosensor platform. *Biosens Bioelectron* 2021;171:112679.
 34. Kukar M, Gunčar G, Vovko T, et al. COVID-19 diagnosis by routine blood tests using machine learning. *Sci Rep* 2021;11:10738.
 35. Wang B, Jin S, Yan Q, et al. AI-assisted CT imaging analysis for COVID-19 screening: Building and deploying a medical AI system. *Appl Soft Comput* 2021;98:106897.
 36. Xu X, Jiang X, Ma C, et al. A Deep Learning System to Screen Novel Coronavirus Disease 2019 Pneumonia. *Engineering (Beijing)* 2020;6:1122-9.
 37. Narin A, Kaya C, Pamuk Z. Automatic detection of coronavirus disease (COVID-19) using X-ray images and deep convolutional neural networks. *Pattern Anal Appl* 2021. [Epub ahead of print]. doi: 10.1007/s10044-021-00984-y.
 38. Zhang J, Xie Y, Li Y, et al. Covid-19 screening on chest X-ray images using deep learning based anomaly detection. *arXiv* 2020:2003.12338.
 39. Arpacı I, Huang S, Al-Emran M, et al. Predicting the COVID-19 infection with fourteen clinical features using machine learning classification algorithms. *Multimed Tools Appl* 2021. [Epub ahead of print]. doi: 10.1007/s11042-020-10340-7.
 40. Villarreal-González R, Acosta-Hoyos AJ, Garzon-Ochoa JA, et al. Anomaly Identification during Polymerase Chain Reaction for Detecting SARS-CoV-2 Using Artificial Intelligence Trained from Simulated Data. *Molecules* 2020;26:20.
 41. Ai T, Yang Z, Hou H, et al. Correlation of Chest CT and RT-PCR Testing for Coronavirus Disease 2019 (COVID-19) in China: A Report of 1014 Cases. *Radiology* 2020;296:E32-40.
 42. Özkaya U, Öztürk Ş, Barstugan M. Coronavirus (COVID-19) classification using deep features fusion and ranking technique. In: Hassanien AE, Dey N, Elghamrawy S. editors. *Big data analytics and artificial intelligence against COVID-19: innovation vision and approach*. Cham: Springer, 2020:281-95.
 43. Wang L, Lin ZQ, Wong A. COVID-Net: a tailored deep convolutional neural network design for detection of COVID-19 cases from chest X-ray images. *Sci Rep* 2020;10:19549.
 44. Gupta A, Anjum, Gupta S, et al. InstaCovNet-19: A deep learning classification model for the detection of COVID-19 patients using Chest X-ray. *Appl Soft Comput* 2021;99:106859.
 45. Paules CI, Marston HD, Fauci AS. Coronavirus Infections-More Than Just the Common Cold. *JAMA* 2020;323:707-8.
 46. Wang Y, Hu M, Li Q, et al. Abnormal respiratory patterns classifier may contribute to large-scale screening of people infected with COVID-19 in an accurate and unobtrusive manner. *arXiv* 2020:2002.05534.
 47. Mouawad P, Dubnov T, Dubnov S. Robust Detection of COVID-19 in Cough Sounds: Using Recurrence Dynamics and Variable Markov Model. *SN Comput Sci* 2021;2:34.
 48. Di Castelnuovo A, Bonaccio M, Costanzo S, et al. Common cardiovascular risk factors and in-hospital mortality in 3,894 patients with COVID-19: survival analysis and machine learning-based findings from the multicentre Italian CORIST Study. *Nutr Metab Cardiovasc Dis* 2020;30:1899-913.
 49. Hu ZJ, Xu J, Yin JM, et al. Lower Circulating Interferon-

- Gamma Is a Risk Factor for Lung Fibrosis in COVID-19 Patients. *Front Immunol* 2020;11:585647.
50. Chassagnon G, Vakalopoulou M, Battistella E, et al. AI-driven quantification, staging and outcome prediction of COVID-19 pneumonia. *Med Image Anal* 2021;67:101860.
 51. Li Y, Shang K, Bian W, et al. Prediction of disease progression in patients with COVID-19 by artificial intelligence assisted lesion quantification. *Sci Rep* 2020;10:22083.
 52. Yang R, Li X, Liu H, et al. Chest CT Severity Score: An Imaging Tool for Assessing Severe COVID-19. *Radiol Cardiothorac Imaging* 2020;2:e200047.
 53. Yan L, Zhang H, Goncalves J, et al. A machine learning-based model for survival prediction in patients with severe COVID-19 infection. *medRxiv* 2020. doi: 10.1101/2020.02.27.20028027.
 54. Pfab J, Phan NM, Si D. DeepTracer for fast de novo cryo-EM protein structure modeling and special studies on CoV-related complexes. *Proc Natl Acad Sci U S A* 2021;118:e2017525118.
 55. Magar R, Yadav P, Barati Farimani A. Potential neutralizing antibodies discovered for novel corona virus using machine learning. *Sci Rep* 2021;11:5261.
 56. Beck BR, Shin B, Choi Y, et al. Predicting commercially available antiviral drugs that may act on the novel coronavirus (SARS-CoV-2) through a drug-target interaction deep learning model. *Comput Struct Biotechnol J* 2020;18:784-90.
 57. Zeng X, Song X, Ma T, et al. Repurpose Open Data to Discover Therapeutics for COVID-19 Using Deep Learning. *J Proteome Res* 2020;19:4624-36.
 58. Erlina L, Paramita RI, Kusuma WA, et al. Virtual screening on Indonesian herbal compounds as COVID-19 supportive therapy: machine learning and pharmacophore modeling approaches. 2020. Available online: https://assets.researchsquare.com/files/rs-29119/v2_stamped.pdf?c=1603225813
 59. Ong E, Wong MU, Huffman A, et al. COVID-19 Coronavirus Vaccine Design Using Reverse Vaccinology and Machine Learning. *Front Immunol* 2020;11:1581.
 60. Wan H, Aravamuthan V, Pearlstein RA. Probing the Dynamic Structure-Function and Structure-Free Energy Relationships of the Coronavirus Main Protease with Biodynamics Theory. *ACS Pharmacol Transl Sci* 2020;3:1111-43.
 61. Paul AS, Islam R, Parves MR, et al. Cysteine focused covalent inhibitors against the main protease of SARS-CoV-2. *J Biomol Struct Dyn* 2020. [Epub ahead of print]. doi: 10.1080/07391102.2020.1831610.
 62. Ke YY, Peng TT, Yeh TK, et al. Artificial intelligence approach fighting COVID-19 with repurposing drugs. *Biomed J* 2020;43:355-62.
 63. Ge Y, Tian T, Huang S, et al. A data-driven drug repositioning framework discovered a potential therapeutic agent targeting COVID-19. *bioRxiv* 2020. doi: 10.1101/2020.03.11.986836.
 64. Morselli Gysi D, do Valle Í, Zitnik M, et al. Network medicine framework for identifying drug-repurposing opportunities for COVID-19. *Proc Natl Acad Sci U S A* 2021;118:e2025581118.
 65. Li L, Li JW, Zhang YM, et al. Material basis and action mechanism of Guiqi Baizhu Fang in the treatment of COVID-19 based on molecular docking prediction for target ACE2 and IL-6R. *Journal of Gansu University of Chinese Medicine* 2020;37:1-9.
 66. Wang J, Mei ZN, Yuan QY, et al. Study on the virtual screening of potential active constituents against coronavirus disease 2019 (COVID-19) of Dao-di herbs in Hubei Province. *Chinese Journal of Hospital Pharmacy* 2020;40:1799-804.
 67. Liu YR, Tang ZS, Wang M, et al. Potential SARS-CoV-2 3CL protease inhibitors selection from TCMSP platform by molecular docking approach. *Chinese Traditional and Herbal Drugs* 2020;51:1694-703.
 68. Enayatkhani M, Hasaniazad M, Faezi S, et al. Reverse vaccinology approach to design a novel multi-epitope vaccine candidate against COVID-19: an in silico study. *J Biomol Struct Dyn* 2021;39:2857-72.
 69. Zawawi A, Forman R, Smith H, et al. In silico design of a T-cell epitope vaccine candidate for parasitic helminth infection. *PLoS Pathog* 2020;16:e1008243.
 70. Senior AW, Evans R, Jumper J, et al. Improved protein structure prediction using potentials from deep learning. *Nature* 2020;577:706-10.
 71. Malone B, Simovski B, Moliné C, et al. Artificial intelligence predicts the immunogenic landscape of SARS-CoV-2 leading to universal blueprints for vaccine designs. *Sci Rep* 2020;10:22375.
 72. Figueroa CA, Aguilera A. The Need for a Mental Health Technology Revolution in the COVID-19 Pandemic. *Front Psychiatry* 2020;11:523.
 73. Fernández RS, Crivelli L, Guimet NM, et al. Psychological distress associated with COVID-19 quarantine: Latent profile analysis, outcome prediction and mediation analysis. *J Affect Disord* 2020;277:75-84.
 74. Choi S, Hong JY, Kim YJ, et al. Predicting Psychological

- Distress Amid the COVID-19 Pandemic by Machine Learning: Discrimination and Coping Mechanisms of Korean Immigrants in the U.S. *Int J Environ Res Public Health* 2020;17:6057.
75. Wang C, Zhao H, Zhang H. Chinese College Students Have Higher Anxiety in New Semester of Online Learning During COVID-19: A Machine Learning Approach. *Front Psychol* 2020;11:587413.
 76. Jha IP, Awasthi R, Kumar A, et al. Explainable-machine-learning to discover drivers and to predict mental illness during COVID-19. *medRxiv* 2020. doi: 10.1101/2020.07.19.20157164.
 77. Kang Y, Kim YJ, Park S, et al. Development and operation of a digital platform for sharing pathology image data. *BMC Med Inform Decis Mak* 2021;21:114.
 78. Ćosić K, Popović S, Šarlija M, et al. Impact of Human Disasters and COVID-19 Pandemic on Mental Health: Potential of Digital Psychiatry. *Psychiatr Danub* 2020;32:25-31.
 79. Chen E, Lerman K, Ferrara E. Tracking Social Media Discourse About the COVID-19 Pandemic: Development of a Public Coronavirus Twitter Data Set. *JMIR Public Health Surveill* 2020;6:e19273.
 80. Li C, Chen LJ, Chen X, et al. Retrospective analysis of the possibility of predicting the COVID-19 outbreak from Internet searches and social media data, China, 2020. *Euro Surveill* 2020;25:2000199.
 81. Hussain W. Role of social media in COVID-19 pandemic. *The International Journal of Frontier Sciences* 2020;4:59-60.
 82. Kumar R, Wang W, Kumar J, et al. An Integration of blockchain and AI for secure data sharing and detection of CT images for the hospitals. *Comput Med Imaging Graph* 2021;87:101812.
 83. Ting DSW, Carin L, Dzau V, et al. Digital technology and COVID-19. *Nat Med* 2020;26:459-61.
 84. Wu J, Wang J, Nicholas S, et al. Application of Big Data Technology for COVID-19 Prevention and Control in China: Lessons and Recommendations. *J Med Internet Res* 2020;22:e21980.
 85. Secundo G, Shams SMR, Nucci F. Digital technologies and collective intelligence for healthcare ecosystem: optimizing Internet of Things adoption for pandemic management. *J Bus Res* 2021;131:563-72.
 86. Rao ASRS, Vazquez JA. Better Hybrid Systems for Disease Detections and Early Predictions. *Clin Infect Dis* 2021. [Epub ahead of print]. doi: 10.1093/cid/ciab489.

Cite this article as: Chang Z, Zhan Z, Zhao Z, You Z, Liu Y, Yan Z, Fu Y, Liang W, Zhao L. Application of artificial intelligence in COVID-19 medical area: a systematic review. *J Thorac Dis* 2021;13(12):7034-7053. doi: 10.21037/jtd-21-747