Since January 2020 Elsevier has created a COVID-19 resource centre with free information in English and Mandarin on the novel coronavirus COVID-19. The COVID-19 resource centre is hosted on Elsevier Connect, the company's public news and information website.

Journal of Infection and Public Health 15 (2022) 289–296

Contents lists available at ScienceDirect

Journal of Infection and Public Health

journal homepage: www.elsevier.com/locate/jiph

Review

Covid-19 and Artificial Intelligence: Genome sequencing, drug development and vaccine discovery

Sali Abubaker Bagabir[a], Nahla Khamis Ibrahim[b,c,*], Hala Abubaker Bagabir[d], Raghdah Hashem Ateeq[e]

[a] Medical Laboratory Technology Department, Faculty of Applied Medical Sciences, Jazan University, Jazan, Saudi Arabia
[b] Community Medicine Department, Faculty of Medicine, King Abdulaziz University, Jeddah, Saudi Arabia
[c] Epidemiology Department, High Institute of Public Health, Alexandria University, Alexandria, Egypt
[d] Medical Physiology Department, Faculty of Medicine, King Abdulaziz University, Rabigh, Saudi Arabia
[e] Faculty of Medicine, King Abdulaziz University, Jeddah, Saudi Arabia

ARTICLE INFO

Article history:
Received 23 September 2021
Received in revised form 11 January 2022
Accepted 12 January 2022

Keywords:
COVID-19
Artificial Intelligence
Genome sequencing
Drugs
Vaccines
Challenges

ABSTRACT

Objectives: To clarify the work done by using AI for identifying the genomic sequences, development of drugs and vaccines for COVID-19 and to recognize the advantages and challenges of using such technology.
Methods: A non-systematic review was done. All articles published on Pub-Med, Medline, Google, and Google Scholar on AI or digital health regarding genomic sequencing, drug development, and vaccines of COVID-19 were scrutinized and summarized.
Results: The sequence of SARS- CoV-2 was identified with the help of AI. It can help also in the prompt identification of variants of concern (VOC) as delta strains and Omicron. Furthermore, there are many drugs applied with the help of AI. These drugs included Atazanavir, Remdesivir, Efavirenz, Ritonavir, and Dolutegravir, PARP1 inhibitors (Olaparib and CVL218 which is Mefuparib hydrochloride), Abacavir, Roflumilast, Almitrine, and Mesylate. Many vaccines were developed utilizing the new technology of bioinformatics, databases, immune-informatics, machine learning, and reverse vaccinology to the whole SARS-CoV-2 proteomes or the structural proteins. Examples of these vaccines are the messenger RNA and viral vector vaccines. AI provides cost-saving and agility. However, the challenges of its usage are the difficulty of collecting data, the internal and external validation, ethical consideration, therapeutic effect, and the time needed for clinical trials after drug approval. Moreover, there is a common problem in the deep learning (DL) model which is the shortage of interpretability.
Conclusion: The growth of AI techniques in health care opened a broad gate for discovering the genomic sequences of the COVID-19 virus and the VOC. AI helps also in the development of vaccines and drugs (including drug repurposing) to obtain potential preventive and therapeutic agents for controlling the COVID-19 pandemic.

Contents

* Correspondence to: Community Medicine Department, KAU, Jeddah, Saudi Arabia.
E-mail address: nibrahim@kau.edu.sa (N.K. Ibrahim).

https://doi.org/10.1016/j.jiph.2022.01.011
1876-0341/

Introduction

Severe Acute Respiratory Syndrome Coronavirus 2 (SARS-CoV-2) causing COVID-19 is one of the most important devastating emerging infectious diseases. COVID-19 pandemic leaves enormous overwhelming loads on wellbeing, survives, economy, with limitation of the globe's social activities, closing borders, banning travel, shutting schools and businesses, applying strict isolation and quarantines, and others [1,2].

Artificial Intelligence (AI) helps on identifying the genomic sequencing of the SARS-CoV-2 virus. Furthermore, there are many strains of SARS-CoV-2. The latest Omicron (B.1.1.529) variants of concern (VOC) has been accompanied by panic responses around the world due to its contagious and vaccine escape mutations. The essential infectivity and antibody resistance of the SARS-CoV-2 variant is determined by its mutations on the spike (S) protein receptor-binding domain (RBD) [3]. AI can help in the rapid discovery of new variant strains. For example, a complete experimental evaluation of the Omicron variant might take weeks or even months, if AI was not applied. Omicron variant mutations are widely distributed on multiple proteins of SARS-CoV-2 such as nonstructural protein (NSP) as (NSP3, NSP4, NSP5, NSP6, NSP12, NSP14, S, envelope, membrane, and nucleocapsid proteins [3].

For the quickly increasing pandemic as COVID-19, AI can also make vaccines and drugs more efficient [4]. The morbidity and death results from COVID-19 can be declined through a rapid intervention, which is possible through an accurate and prompt prediction of disease advance. Progressions in AI algorithms assist to analyze a countless volume of data and create significant predictions, decisions, and automation [5]. Nowadays, several target drugs have been identified and listed to be potential inhibitors of SARS-CoV-2 [6,7]. For the development of drugs and vaccines for COVID-19, AI can improve the identification of genetic materials [6]. AI technology has become a digital arming in the development of new drugs, vaccines, diagnostic methods, and forecasting programs [7]. COVID-19 is the most important global health problem nowadays. However, scanty studies were conducted for reviewing the research done on use of AI for helping in the controlling COVID-19 pandemic. In consequence, a review on AI application for identifying the genomic sequence, development of drugs and vaccines for COVID-19 is essential.

The objective of the current study was to clarify the work done by using AI for identifying the genomic sequences, development of drugs and vaccines for COVID-19 and to recognize the advantages and challenges of using such technology.

Methods

A non-systematic review was done. All articles published on PubMed, Medline, Google, Google scholar on the use of AI and/or digital health for identifying the genomic sequencing, drug development, and vaccine discovery for SARS-CoV-2 were scrutinized and summarized. Results of the advantages, challenges and the limitations of using such methods were also summarized.

Results

A total of 38 papers fitting the objectives of the study were selected. The results are presented in the following sections.

Journey of COVID-19 genomic sequencing through the help of AI

Regarding the genomic sequencing of the SARS-CoV-2 virus, Lopez-Rincon et al. [8] used the deep learning (DL) method for DNA sequences classification using Convolutional Neural Networks (CNN). The training of CNN has been used to separate different viral genome strains from the family of Coronaviruses. Then, AI techniques have been applied to find representative sequences. In this method the network was utilized to identify SARS-CoV-2 genomic sequence. The discovered sequences have been validated using different experiments and different repositories (Datasets) such as the National Center for Biotechnology Information (NCBI), the Global Initiative on Sharing All Influenza Data (GISAID), and the National Genomics Data Center (NGDE). This is done to prove the ability to separate the complementary DNA (cDNA), which is the DNA synthesized from a single-stranded RNA template in a reaction catalyzed by the enzyme reverse transcriptase. The use of the cDNA sequence of SARS-CoV-2 from different virus strains was done for being near to the perfection accuracy, and to create an innovative set of viral sequence features. Twelve sequences of 21-base-pairs (bps) of SARS-CoV-2 were specified using this procedure. One of these feature sequences is nominated to create a primer set. The primer sets are tested *versus* the state-of-the-art primer sets, synthesized, and hence used for laboratory testing. It is observed that this primer set has a high specificity for the detection of SARS-CoV-2. In addition, it was found that this procedure evaluates SARS-CoV-2 with an accuracy of more than 99%. However, in Quantitative reverse transcription PCR (RT-qPCR), the accuracy is around 70%. Meanwhile, the CT scan coupled with DL to reach an accuracy of 83% [8]. In 2021, a viral reverse engineering design was done to find out the similarity in viral protein & genomic sequences, and the additional mutations were extracted depending on the phylogenetic tree for capturing the evolutionary behavior [5].

Concerning the Omicron VOC, an AI model has been trained with tens of thousands of experimental data points data on SARS-CoV-2 [3]. An efficient and reliable in-silico analysis is imperative and valued for such an urgent situation. A comprehensive topology-based AI model called TopNetmAb was used to predict the binding free energy (BFE) changes of S and ACE2/antibody complexes induced by mutations on the spike RBD. The positive BFE change induced by a specific RBD mutation indicates its potential ability to strengthen the binding of an S protein-ACE2/antibody complex, while a negative BFE change suggests a likely capacity to reduce the binding strength of an S protein-ACE2/antibody complex. Since the S protein, particularly its RBD, plays a vital role in viral infection, it has been a key target of vaccines and antibody drugs. The study of Omicron's 15 RBD mutations, and based on a well-tested and experimentally confirmed DL model trained with tens of thousands of experimental data, the impacts of Omicron's RBD mutations on its infectivity were investigated. Results found that Omicron is about ten times more infectious than the original virus or about twice as infectious as the Delta variant. Using the structures of 132 known antibody-RBD complexes, it was found that Omicron's vaccine-escape capability is nearly twice as high as that of the Delta variant. Omicron may significantly reduce the efficacy of the Eli Lilly antibody cocktail. Omicron RBD mutations may also compromise monoclonal antibodies (mAbs) [3].

Journey of COVID-19 drugs discovering through the help of AI

Several studies report the usage of AI to generate modern drug compounds. Zhavoronkov et al. [9] used an AI-based drug discovery pipeline to generate a new drug compound that is cost-effective and with time productivity. Those compounds are inhibitors for the SARS-CoV2 3C-like protease ($3CL^{pro}$). Tang et al. [10] developed an advanced deep Q-learning network with the fragment-based drug design (ADQN-FBDD) (a model-free reinforcement learning algorithm) to generate novel lead compounds. Based on their AI methodology, different derivatives were obtained from the generated lead compounds (47 lead compounds). Gao et al. [11] used an AI-based generative network complex to generate 15 potential drugs. When those drugs were compared with HIV inhibitors, it was found that it has enhanced drug properties [11]. Hofmarcher et al. [12] used a deep neural network protocol on three drug discovery databases to generate small compounds that are SARS-CoV-2 inhibitors. Firstly, they screened around one billion compounds from a database and then reduced them to 30 thousand compounds. Those compounds are accessible in the ZINC Database (collection of commercially available chemical compounds prepared especially for virtual screening). Furthermore, those 30,000 compounds are available in the library on: https://github.com/ML-jku/sars-cov-inhibitors-chemai.

Zhang et al. created a DL protocol and used four chemical compounds and tripeptides databases to identify potential drugs for COVID-19. A list of chemical ligands and peptide drugs was provided [13].

It is generally believed that the development of a new drug for the therapy of COVID-19 takes a long time. It is commonly considered that the therapeutic shortcut is to use the drug repurposing technique to find new uses of the already existing drugs [13]. Concerning drug repurposing, several studies used such a strategy based on DL algorithms. Beck et al. used the DL-based architecture to estimate drug-target interactions that could disturb the component of SARS-CoV-2 using the model called 'Molecule Transformer-Drug Target Interaction (MT-DTI). Based on their procedure, a list of antiviral drugs was identified (Atazanavir, Remdesivir, Efavirenz, Ritonavir, and Dolutegravir). The best chemical compound is Atazanavir, which is used to prevent HIV/AIDS. Compared with other listed drugs, Atazanavir is shown to be effective for the COVID-19 treatment as it has a high ability to bind to six proteins of SARS-CoV-2; including the replication component of SARS-CoV-2, $3CL^{pro}$, and RNA-dependent RNA polymerase (RdRP, RDR). However, the *in vitro* and *in vivo* experimental study is required to validate the findings for safety and efficacy [14]. Meanwhile, in another study, Ge et al. used a drug repurposing strategy based on machine learning (ML) and statistical analysis to retrieve data related to coronavirus and the drug candidates followed by experimental validation. Based on the computational framework, a list of drug candidates was observed. It is discovered that poly-ADP-ribose polymerase 1 (PARP1) inhibitors showed antiviral activities against SARS-CoV-2. Two PARP1 inhibitors were selected, Olaparib, which is Food and Drug Administration (FDA) approved drug, and CVL218, which is Mefuparib hydrochloride and is still under clinical trial. Experimental studies, both *in vitro* and *in vivo*, have been conducted on these two drugs and several other tested drugs. Results revealed that CVL218 is a potential and effective drug that can be used for COVID-19 treatment. *In vitro* studies revealed that 3 μM and 30 μM doses of CVL218 showed successful inhibitory activity against SARS-CoV-2 replication by 35.16% and 99.68%, respectively. The results showed no evidence of a cytopathogenic effect. Furthermore, pharmacokinetic and toxicokinetic studies, *in vivo*, indicated that the high concentration of CVL218 showed no sign of toxicity [15].

Previous conducted studies illustrated that patients with SARS-CoV-2 have a significant increased level of interleukin-6 (IL-6), and hence an outrageous stimulated immune response [16–18]. According to Ge et al. [15] the *in vitro* studies done showed that CVL218 has the ability to suppress the production of IL-6 that induced by CpG DNA (regions of DNA) in peripheral blood mononuclear cells. Hence, it was suggested that it might have an anti-inflammatory effect to prevent immunopathology generated by the infection of SARS-CoV-2. In addition, a molecular docking simulation was conducted and showed that CVL218 binds more effectively than Olaparib. Hence, it decreases its RNA binding and inhibits viral replication.

In another study, Hu, and his colleagues used a computational-based methodology for *in silico* screening (that may decrease the time and expense of discovering treatment). They found that the drug abacavir is expected to have a high affinity to bind to numerous proteins of SARS-CoV-2. Abacavir is a nucleoside analog reverse-transcriptase inhibitor (NRTI) which applied for the treatment of HIV/AIDS. It is also discovered that roflumilast and almitrine mesylate are expected to show an inhibitory outcome. Furthermore, according to the procedure of Hu et al. a list of ten target drugs that showed a potential inhibitory effect against SARS-CoV-2 are generated [19]. In March 2020, an integrative, antiviral drug repurposing methodology implementing a systems pharmacology-based network medicine platform was applied for drug repurposing. Sixteen candidate drugs were identified and suggested to be potential drugs that can be used for the treatment of the COVID-19 virus [20].

In July 2020, another study using a network-deep learning methodology in which a medical knowledge graph named CoV-KGE was constructed. It consisted of 15 million edges crosswise different types of relationships between drugs, proteins, diseases, genes, protein expression, and pathways. In addition to the computing resources of Amazon Web Service (AWS), which is a cloud provider where 41 repurposing drugs have been discovered. These drugs were then authenticated *via* three transcriptomic and one proteomics data set in COVID-19 infected human cell-line as well as data from the current clinical trials [21].

A similar study was reported in August 2020. It constructed a comprehensive graph neural network. The authors utilized the developments of the network medicine for defining 81 drug repurposing candidates for the treatment of COVID-19 by using *in vitro* data. They reported that such drugs affected the biological processes targeted by the virus [22].

Traditional Chinese Medicine (TCM) can be considered as a promising treatment for COVID-19. However, it has been reported that there are many side effects. Hence, the ontology-based side-effect prediction framework (OSPF) was developed. This protocol will separate the ingredients of TCM into two main categories; either hot or cold. After that, a database was established to convert each TCM prescription into a vector that contains ingredients dosage according to the two categories and label as safe and unsafe. Artificial Neural Network (ANN)-based DL used this database to evaluate the TCM prescription in which the safety indicator (SI) is given (High to Low). SI is the percentage of the possibility of occurrence of side effects. Then, the list of proposed TCM prescriptions is re-arranged from high to low. It was stated that seven of them have an indicator of more than 0.8 in which will consider in the therapeutics of SARS-CoV-2 [23].

In Egypt, a new framework called DL for prediction drug-target affinities (DeepH-DTA) was proposed for predicting DT binding affinities for heterogeneous drugs. Heterogeneous graph attention (HGAT) model was developed to learn topological information of compound molecules and bidirectional ConvLSTM layers for modeling spatio-sequential information in simplified molecular-input line-entry system (SMILES) sequences of drug data. For protein sequences, a squeezed-excited dense convolutional network for learning hidden representations within amino acid sequences was proposed; while utilizing advanced embedding techniques for encoding both kinds of input sequences. A study using FDA-approved

drugs from the Drug Bank database is performed using DeepH-DTA to predict the affinity scores of drugs against SARS-CoV-2 amino acid sequences. The results showed that the model can predict some of the SARS-Cov-2 inhibitors that have been recently approved in many clinical studies [24].

Currently, one way to predict the treatment of viral diseases is to identify the MicroRNAs (miRNAs). They are non-coding and small RNAs that have an essential role in gene expression regulation. Demirci and Adan [25] used a machine learning-based miRNA prediction methodology for the SARS-CoV-2 for predicting the pre-miRNA and mature miRNA. They searched for the interaction between human genes and viral miRNAs as well as between viral genes and human miRNAs for studying the mechanism of SARS-CoV-2 infection. They concluded that when the level of human miRNAs is increased in a way that targets the structural protein of SARS-CoV-2, the viral entry and replication would be blocked. Meanwhile, when the level of human miRNAs is decreased, the viral entry and replication would be increased, and hence it will be more visible for the human immune system. This could build a new window for the development of new therapeutics. However, it was stated that more *in vivo* and *in vitro* experiments are needed to validate the miRNAs candidate [25].

Journey of COVID-19 vaccine discovery through the help of AI

The types of vaccine discovered for COVID-19 until now are either: the whole virus (inactivated or attenuated), viral vector (replicating and non-replicating), nucleic acid (RNA and DNA), and protein-based (protein subunit, virus-like particle) [26]. The main advantages of AI are rapid and accurate discovery with which it recognizes cases and efficacy in screening for vaccine production. AI modalities are utilized for effective vaccine development and for evaluating the safety of such vaccines. Some of the vaccines that had emergency use approval (EUA) during the pandemic were Messenger RNA (mRNA) BNT162b2 from Pfizer-BioNTech (Pfizer/BioNTech), mRNA-1273 (Moderna), Covishield (Oxford-AstraZeneca), and JNJ-78436735/Ad26. COV2. S (Johnson and Johnson), *etc.* [5].

On the 2nd of March 2020, epitope vaccines (T cell epitope-based peptide vaccines) against SARS-CoV-2 were designed by using protein E as antigenic site [27]. The vaccine was designed by firstly determining the antigenic target in order to determine the peptides in which it performs as ligands. Sophisticated bioinformatics tools and databases were used and determined 20 peptides in protein E as a candidate for SARS-CoV-2. Firstly, the whole genome of SARS-CoV-2 was analyzed by using the comparative genomic method. Then, a comparison between human coronavirus and COVID-19 sequences was carried out using Artemis Comparison Tool (ACT). A high mutation rate was found in the four structural proteins in coronavirus. Those four proteins are the S, Membrane (M), Nucleocapsid (N), and Envelope (E) proteins. These proteins were analyzed by VaxiJen Software (for Prediction of Protective Antigens and Subunit Vaccines) to find the highest possibility of antigenic protein. Protein E was chosen at that time as an immunogenic target. After that, binding affinities to the major histocompatibility complex (MHC) which are the MHC class I and II were evaluated to select the peptides that can be used for vaccine design. Three peptides (SLVKPSFYV and LAILTALRL and YVYSRVKNL) were selected with high potentiality for vaccine design. In this study, a recommendation was stated to validate the designed vaccine for safety and immunogenic report. On the 6th of March 2020, a study used the methodology of immune-informatics and reverse vaccinology to design an epitope-based subunit vaccine against COVID-19. Three vaccines were successfully constructed, meanwhile only one vaccine was designated as the greatest vaccine according to the molecular docking analysis [1].

In early March 2020, a computational methodology was also carried out to identify several epitopes in SARS-CoV-2. The viral genome of SARS-CoV-2 was analyzed for identifying several epitope candidates that can be used for the development of potential vaccines. It has been reported that 405T-cell epitopes with strong MHC-I and II scores and two B-Cell epitopes were identified near spike protein. Moreover, while analyzing the 68 viral genome mutation profiles, it is stated that there is no mutation found near the S protein. Consequently, it is identified as an immunogenic and effective vaccine candidate [28]. Hence, the spike protein has been discovered nowadays as the most intergenic gene used for vaccine discovery.

Ong et al. has predicted a vaccine against COVID-19 using the ML and reverse vaccinology to the whole SARS-CoV-2 proteomes as well as the structural and non-structural proteins. Accordingly, they have found six proteins that are predicted to be adhesions (one structural protein S-protein and five non-structural proteins (nsp3, 3CL-pro, nsp8, nsp9, and nsp10)). In addition, they reported that three of these proteins (S protein, nsp3, and nsp8 proteins) have high protective antigenicity scores. The nsp3 protein has not been tested before, and hence, in their study, they investigated more about this protein using ML methodology. It has been found that nsp3 is more conserved between Middle East Respiratory Syndrome (MERS-CoV), SARS-CoV, and SARS-CoV-2. Adding to that, it is predicted to have promiscuous MHC-I and MHC-II T-cell and B-cell epitopes. Therefore, it is indicated that many of the non-structural proteins can be applied as potential vaccine candidates. Furthermore, they proposed a cocktail vaccine with structural and non-structural proteins; such as the S protein/nsp3 cocktail vaccine in order to cover the different sides of protection, and it would accelerate efficient complementary immune responses [29].

A multi-epitope-based chimeric vaccine design against structural proteins of SARS-CoV-2 was constructed. Rahman et al. used an advanced bioinformatics technique (an immune-informatics methodology) alongside comparative genomics to develop an effective chimeric vaccine and named CoV-RMEN [30]. Susithra Priyadarshni et al. [31] used an in silico approach to design a multi-epitopic vaccine candidate targeting the non-mutational immunogenic regions in envelope protein and surface glycoprotein of SARS-CoV-2.

The emerging of lethal and new VOC of COVID-19, such as the delta variants and Omicron puts the health systems under a burden, and yields hesitancy concerning the upcoming efficacy of the existing vaccines. Adding to prevention, the therapy of COVID-19 is also affected by emergent strains. So, AI modalities have been established as a fundamental method for tracking, and forecasting the mutational sites and for rapid and adequate application of prevention and control measures of COVID-19 [5].

Concerning Omicron and based on 132 three-dimensional (3D) structures of antibody- RBD complexes, it was found that Omicron is maybe twice more likely to escape current vaccines than the Delta variant. The Food and Drug Administration (FDA)-approved mAbs from Eli Lilly may be seriously compromised by this variant. Omicron may also diminish the efficacy of mAbs from Celltrion and Rockefeller University. However, its impact on the Regeneron mAb cocktail appears to be mild [3]. The combined use of both molecular prediction tools and computer simulation in the development or regulatory evaluation of medical intervention is making the difference to better predict the efficacy and safety of new vaccines. An integrated bioinformatics pipeline that merges the prediction power of different software that act at different scales for evaluating the elicited response of the human immune system against every pathogen is proposed. The authors applied a problem-solving protocol to predict the cross-reactivity of pre-existing vaccination interventions against SARS-CoV-2. This study sheds the light on the fact that they can support the '3 Rs' principles (replace, reduce, refine) in drug development and better predict the efficacy of new vaccines. Moreover, in the context of the VOC, the proposed in silico pipeline

Table 1
Characteristics of the included articles on the use of Artificial Intelligence on COVID-19 diagnosis, drug and vaccine discovery.

First author (year) Ref.	Area or specialty of AI application	AI application method/AI MODEL	Clinical benefit	Country
Lopez-Rincon (2021) [8]	Diagnosis	DL method (3D-DL framework) for DNA sequences classification using CNN.	Viral genomic sequence of SARS-Cov2 Twelve sequences of 21-base-pairs (bps) of SARS-CoV-2 were specified. One is nominated to create a primer set. The primer set is used to detect SARS-CoV-2. More than 99% accuracy in viral evaluation.	Netherlands.
Chen et al. (2021) [3]	Diagnosis	TopNetmAb model: comprehensive topology-based AI	Predict the BFE changes of S and ACE2/antibody complexes induced by mutations on the S RBD, of the Omicron variant. Omicron's vaccine-escape capability is about twice as high as that of the Delta variant. Omicron may significantly reduce the efficacy of the Eli Lilly antibody cocktail. It may also compromise mAbs.	USA
Zhavoronkov et al. (2020) [9]	Drug discovery	Generative DL. AI-based drug discovery pipeline.	Generate a new drug compound that is cost-effective and with time productive. They are inhibitors for the SARS-CoV-2 $3CL^{pro}$.	Hong Kong
Tang et al. (2020) [10]	Drug discovery	ADQN-FBDD: Advanced deep Q-learning network with the fragment-based drug design (a model-free reinforcement learning algorithm).	Generate novel lead compounds (47 lead compounds were discovered) targeting SARS-CoV2 3C-like main protease.	China
Gao et al. (2020) [11].	Drug discovery	AI-based generative network complex (GNC)	Generate 15 potential drugs. When those drugs have enhanced drug properties.	USA.
Hofmarcher et al. (2020) [12].	Drug discovery	ChemAI; Deep neural network protocol on three drug discovery databases.	Generate small compounds that are SARS-CoV-2 inhibitors. 30,000 compounds are screened and available in the library (ZINC database).	China
Zhang et al. (2020) [13]	Drug discovery	Dense Fully Convolutional Neural Network (DFCNN)A DL protocol and used four chemical compounds and tripeptides databases to identify potential drugs for COVID-19.	A list of chemical ligands and peptide drugs was provided.	China
Beck et al. (2020) [14]	Drug discovery	Molecule Transformer-Drug Target Interaction (MT-DTI).	Estimate drug-target interactions.A list of antiviral drugs was identified. The best chemical compound is Atazanavir, which is used to prevent HIV/AIDS.	The Republic of Korea.
Ge et al. (2020) [15]	Drug discovery	A drug repurposing strategy based on ML CoV-DTI: a network-based knowledge mining algorithm with the integrative framework involved ML and statistical analysis methods followed by an experimental validation (*in vitro* and *in vivo* experimental study), as well as a molecular docking simulation.	PARP1 inhibitors showed antiviral activities against SARS-CoV-2. Two PARP1 inhibitors are selected; Olaparib and CVL218 CVL218 is a potential and effective drug for COVID-19 treatment. *In vitro* studies, it is stated that 30 μM doses of CVL218 showed successful inhibitory activity against SARS-CoV-2 replication by 99.68%. Moreover, it has the ability to suppress the production of IL-6. *In vivo* pharmacokinetic and toxicokinetic studies, it is found that a high concentration of CVL218 is not toxic.	China
Hu et al. (2020) [19]	Drug discovery	Pre-trained multi-task deep model. A computational-based methodology for in Silico screening.	Decrease time and expense of discovering treatment. Ten target drugs that showed potential inhibitory effects are generated. Abacavir, roflumilast, and almitrine mesylate are expected to show an inhibitory outcome.	Hong Kong
Zhou et al. (2020) [20]	Drug discovery	An integrative, antiviral drug repurposing with a pharmacology-based network medicine platform.	Sixteen candidate drugs were identified and suggested to be potential drugs	USA
Zeng et al. (2020) [21]	Drug discovery	A network-DL methodology with a graph named CoV-KGE. Computing resources of Amazon Web Service (AWS).	A cloud provider with 41 repurposing drugs has been discovered.	China
Gysi et al. (2020) [22]	Drug discovery	A comprehensive graph neural network.	Defining 81 drug repurposing candidates by using *in vitro* data. The drugs affected the biological processes targeted by the virus.	USA

(*continued on next page*)

Table 1 (*continued*)

First author (year) Ref.	Area or specialty of AI application	AI application method/AI MODEL	Clinical benefit	Country
Wang et al. (2020) [23]	Drug discovery	Ontology-based side-effect prediction framework (OSPF). ANN-based DL was used to evaluate the TCM prescription with the SI.	Seven of TCM have high safety indicators (SI of more than 0.8).	China
Abdel-Basset et al. (2020) [24]	Drug discovery	DeepH-DTA: a HGAT model was developed to learn topological information of compound molecules and bidirectional ConvLSTM layers for modeling spatio-sequential information in SMILES sequences of drug data. A squeezed-excited dense convolutional network for learning hidden representations within amino acid sequences.	Predict the affinity scores of drugs against SARS-CoV-2 amino acid sequences. Predict some of the SARS-CoV-2 inhibitors in which they are FDA-approved drugs.	Egypt
Demirci et al. (2020) [25]	Drug discovery	ML-based miRNA prediction analysis. Computational analysis of miRNA-mediated interactions.	Study the mechanism of SARS-CoV-2 infection. When human miRNAs are increased, the viral entry and replication would be blocked. An opportunity for the development of new therapeutics.	Turkey
Abdelmageed et al. (2020) [27]	Vaccine discovery	Bioinformatics tools and databases (comparative genomic method, ACT, and VaxiJen Software).	Epitope vaccines were designed by using protein E as an antigenic site.	Sudan
Sarkar et al. (2020) [1]	Vaccine discovery	Immune-informatics, reverse vaccinology, and molecular docking analysis.	Three epitope-based subunit vaccines were designated. only one was reported to be as the greatest vaccine	Bangladesh
Fast et al. (2020) [28]	Vaccine discovery	Computational methodology.	Identify several epitopes in SARS-CoV-2 for the development of potential vaccines. S protein was identified as an immunogenic and effective vaccine candidate.	USA
Ong et al. (2020) [29]	Vaccine discovery	ML and reverse vaccinology	It is indicated many non-structural proteins can be applied as potential vaccine candidates. It is proposed to use a cocktail vaccine with structural and non-structural proteins in which would accelerate efficient complementary immune responses.	USA
Rahman et al. (2020) [30]	Vaccine discovery	Immuno-informatics approach along with comparative genomics. In silico approach.	A multi-epitope-based chimeric peptide vaccine is designed against S, M, and E proteins and named CoV-RMEN.	Bangladesh.
Susithra Priyadarshni et al. (2021) [31]	Vaccine discovery	In silico approach.Molecular docking analysis	Design a multi-epitopic vaccine candidate targeting the non-mutational immunogenic regions in envelope protein and surface glycoprotein of SARS-CoV-2	USA
Russo et al. (2021) [32]	Vaccine discovery	An integrated bioinformatics pipeline that merges the prediction power of different software (in silico pipeline)	Predict the cross-reactivity of pre-existing vaccination interventions against SARS-CoV-2. The proposed in silico pipeline can be applied to predict the potential cross-reactive immunity induced by existing vaccinations against SARS-CoV-2 new emerging variants.The method can speed up the development of vaccines tailored to the emerging antigenic variants as the Omicron.	Italy

can be applied to predict the potential cross-reactive immunity induced by existing vaccinations against SARS-CoV-2 new emerging variants (*e.g.* B.1.1.7, B.1.351, and other lineages). Together with the possibility to use this problem-solving protocol to estimate the degree of the induced immune response by completely newly developed vaccines, it can speed up the development of vaccines tailored to the emerging antigenic variants [32]. Table 1.

Discussion

Many types of AI were used in the journey of genomic sequencing discovery, development of drugs and vaccines for COVID-19. The AI types used in the current work are ML, DL, and ANN [5].

The AI approaches open a new gate for all scientists in a different discipline to retrieve a huge amount of useful biomedical data in a short time to identify genomic sequence, design and discover a vaccine or a new therapy for emerging diseases. In addition, AI is a core of advanced technologies where the discovered therapy would be available for a clinical trial, and when it is approved it uses for health care [33]. AI provides cost-saving and agility methods in Covid-19 when it is approved it can be used for health care [34]. Genomic surveillance and AI are the main weapons for the early detection of new VOC of SARS-CoV-2. Detecting VOC and developing an evidence-based and timely public health response to mitigate the spread of the new variants, requires a robust genomic surveillance program. Monitoring is also required for assessing the effectiveness of vaccines against the circulating virus. Data need to be publicly available to drive real-time decision-making, enhance transparency, and fuel relevant research by national and international institutions [35]. However, other factors should be taken into account as challenges and limitations for the use of AI. Some of these challenges are represented by the forecast of the side effects, the rate of drug production, the therapeutic effect and the drug costs [19]. DL is used to classify virus species dividing the sequences into a fixed-length that varies from 300 bps to 3000 bps cannot be filled completely [8]. The data should also be reliable and of high quality to ensure accurate result prediction. During the occurrence of disease, the amount of data about it is not always enough to o retrieve information and predict algorithms (AI-based algorithms) needed to implement further analysis about the disease or to discover specific vaccines or drugs. Takes also further time, hence the time to discover the new vaccines and drugs is very crucial. The external validation should be performed in different geographical regions. Hence, the algorithms will be more accurate for those specific populations rather than in other geographical regions [36–38].

Conclusion

The growth of AI techniques in health care opened a broad gate for discovering the genomic sequences of the COVID-19 virus and the VOC. AI helps also in the development of vaccines and drugs (including drug repurposing) to obtain potential preventive and therapeutic agents for controlling the COVID-19 pandemic. The AI types used in the current study are ML, DL, and ANN [5] are the commonest used methods. AI provides cost-saving and agility methods for prevention and control of Covid-19. However, the challenges of its usage are the difficulty of collecting data, the internal and external validation, ethical consideration, therapeutic effect, and the time needed for clinical trials after drug approval.

Funding

No funding Sources.

Ethical approval

Not required.

Competing interests

None declared.

Acknowledgements

There is no fund.

References

[1] Sarkar B, Ullah MA, Johora FT, Taniya MA, Araf Y. The essential facts of Wuhan novel coronavirus outbreak in China and epitope-based vaccine designing against 2019-nCoV. BioRxiv 2020. https://doi.org/10.1101/2020.02.05.935072
[2] Ibrahim NK. Epidemiologic surveillance for controlling Covid-19 pandemic: types, challenges and implications. J Infect Public Health 2020;13:1630–8. https://doi.org/10.1016/j.jiph.2020.07.019
[3] Chen J, Wang R, Gilby NB, Wei GW. Omicron (B.1.1.529): infectivity, vaccine breakthrough, and antibody resistance. ArXiv Prepr 2021. [arXiv:2112.01318v1].
[4] Naseem M, Akhund R, Arshad H, Ibrahim MT. 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.
[5] Arora G, Joshi J, Mandal RS, Shrivastava N, Virmani R, Sethi T. Artificial intelligence in surveillance, diagnosis, drug discovery and vaccine development against COVID-19. Pathogens 2021;10(8). https://doi.org/10.3390/pathogens10081048
[6] Jawaid Akhtar M. COVID19 inhibitors: a prospective therapeutics. Bioorg Chem 2020;101:104027https://doi.org/10.1016/j.bioorg.2020.104027
[7] Yanisky-Ravid S, Jin R. Summoning a new artificial intelligence patent model: in the age of pandemic. Ssrn 2020;1–49:3619069 https://doi.org/10.2139/ssrn.3619069
[8] Lopez-Rincon A, Tonda A, Mendoza-Maldonado L, Mulders DGJC, Molenkamp R, Perez-Romero CA, et al. Classification and specific primer design for accurate detection of sars-cov-2 using deep learning. Sci Rep 2021;11(1):1–11.
[9] Zhavoronkov A, Aladinskiy V, Zhebrak A, Zagribelnyy B, Terentiev V, Bezrukov DS, et al. Potential COVID-2019 3C-like protease inhibitors designed using generative deep learning approaches. Insilico Med Hong Kong Ltd A 2020;307:E1.
[10] Tang B, He F, Liu D, Fang M, Wu Z, Xu D. AI-aided design of novel targeted covalent inhibitors against SARS-CoV-2. BioRxiv 2020. https://doi.org/10.1101/2020.03.03.972133
[11] Gao K, Nguyen DD, Wang R, Wei GW. Machine intelligence design of 2019-nCoV drugs. bioRxiv 2020. https://doi.org/10.1101/2020.01.30.927889
[12] Hofmarcher M, Mayr A, Rumetshofer E, Ruch P, Renz P, Schimunek J, et al. Large-scale ligand-based virtual screening for SARS-CoV-2 inhibitors using deep neural networks. Tech Rep 2020 [Available at: SSRN 3561442].
[13] Zhang H, Saravanan KM, Yang Y, Hossain MT, Li J, Ren X, et al. Deep learning based drug screening for novel coronavirus 2019-nCov. Interdiscip Sci: Comput Life Sci 2020;12:368–76.
[14] Beck BR, Shin B, Choi Y, Park S, Kang K. 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. https://doi.org/10.1016/j.csbj.2020.03.025
[15] Ge Y, Tian T, Huang S, Wan F, Li J, Li S, et al. A data-driven drug repositioning framework discovered a potential therapeutic agent targeting COVID-19. Signal Transduct Target Ther 2021;6:165–70.
[16] Zhou Y, Fu B, Zheng X, Wang D, Zhao C, Qi Y, et al. Aberrant pathogenic GM-CSF+ T cells and inflammatory CD14+ CD16+ monocytes in severe pulmonary syndrome patients of a new coronavirus. BioRxiv 2020. https://doi.org/10.1101/2020.02.12.945576
[17] Diao B, Wang C, Tan Y, Chen X, Liu Y, Ning L, et al. Reduction and functional exhaustion of T cells in patients with coronavirus disease 2019 (COVID-19). Front Immunol 2020;11:827. https://doi.org/10.3389/fimmu.2020.00827
[18] Li B, Feng F, Yang G, Liu A, Yang N, Jiang Q, et al. Immunoglobulin G/M and cytokines detections in continuous sera from patients with novel coronaviruses (2019-nCoV) infection. Available at SSRN: ⟨https://ssrn.com/abstract=3543609⟩. ⟨http://dx.doi.org/10.2139/ssrn.3543609⟩.
[19] Hu F, Jiang J, Yin P. Prediction of potential commercially inhibitors against SARS-CoV-2 by multi-task deep model. arXiv Prepr arXiv 2020;200300728.
[20] Zhou Y, Hou Y, Shen J, Huang Y, Martin W, Cheng F. Network-based drug repurposing for novel coronavirus 2019-nCoV/SARS-CoV-2. Cell Discov 2020;6:14. https://doi.org/10.1038/s41421-020-0153-3
[21] Zeng X, Song X, Ma T, Pan X, Zhou Y, Hou Y, et al. Repurpose open data to discover therapeutics for COVID-19 using deep learning. J Proteome Res 2020;19(11):4624–36.
[22] Gysi DM, Do Valle Í, Zitnik M, Ameli A, Gan X, Varol O, et al. Network medicine framework for identifying drug repurposing opportunities for COVID-19. ArXiv 2020.
[23] Wang Z, Li L, Yan J, Yao Y . Evaluating the traditional Chinese medicine (TCM) officially recommended in China for COVID-19 using ontology-based side-effect prediction framework (OSPF) and deep learning; 2020.
[24] Abdel-Basset M, Hawash H, Elhoseny M, Chakrabortty RK, Ryan M. DeepH-DTA: deep learning for predicting drug-target interactions: a case study of COVID-19 drug repurposing. IEEE Access 2020;8:170433–51. https://doi.org/10.1109/access.2020.3024238

[25] Demirci MDS, Adan A. Computational analysis of microRNA-mediated interactions in SARS-CoV-2 infection. PeerJ 2020;8:e9369.
[26] Velikova T, Georgiev T. SARS-CoV-2 vaccines and autoimmune diseases amidst the COVID-19 crisis. Rheuma Int 2021;41(3):509–18. https://doi.org/10.1007/s00296-021-04792-9
[27] Abdelmageed MI, Abdelmoneim AH, Mustafa MI, Elfadol NM, Murshed NS, Shantier SW, et al. Design of multi epitope-based peptide vaccine against E protein of human 2019-nCoV: an immunoinformatics approach. bioRxiv 2020:11. [Preprint posted online February].
[28] Fast E, Altman RB, Chen B. Potential T-cell and B-cell epitopes of 2019-nCoV. bioRxiv 2020. https://doi.org/10.1101/2020.02.19.955484. [2020.02.19.955484].
[29] Ong E, Wong MU, Huffman A, He Y. COVID-19 coronavirus vaccine design using reverse vaccinology and machine learning. Front Immunol 2020;11:1581.
[30] Rahman MS, Hoque MN, Islam MR, Akter S, Rubayet-Ul-Alam ASM, Siddique MA, et al. Epitope-based chimeric peptide vaccine design against S, M and E proteins of SARS-CoV-2 etiologic agent of global pandemic COVID-19: an in silico approach. BioRxiv 2020. [03. 30.015164. 2020].
[31] Susithra Priyadarshni M, Isaac Kirubakaran S, Harish MC. In silico approach to design a multi-epitopic vaccine candidate targeting the non-mutational immunogenic regions in envelope protein and surface glycoprotein of SARS-CoV-2. J Biomol Struct Dyn 2021:1–16. https://doi.org/10.1080/07391102.2021.1977702
[32] Russo G, Di Salvatore V, Sgroi G, Parasiliti Palumbo GA, Reche PA, Pappalardo F. A multi-step and multi-scale bioinformatic protocol to investigate potential SARS-CoV-2 vaccine targets. Brief Bioinform 2021. https://doi.org/10.1093/bib/bbab403
[33] Zhou Y, Wang F, Tang J, Nussinov R, Cheng F. Artificial intelligence in COVID-19 drug repurposing. Lancet Digit Health 2020;2(12):e667–76. https://doi.org/10.1016/S2589-7500(20)30192-8
[34] Klumpp M, Hintze M, Immonen M, Ródenas-Rigla F, Pilati F, Aparicio-Martínez F, et al. Artificial intelligence for hospital health care: application cases and answers to challenges in European hospitals. Healthcare 2021;9. https://doi.org/10.3390/healthcare9080961
[35] Hasan MM, Rocha ICN, Ramos KG, Cedeño TDD, Dos Santos Costa AC, Tsagkaris C, et al. Emergence of highly infectious SARS-CoV-2 variants in Bangladesh: the need for systematic genetic surveillance as a public health strategy. Trop Med Health 2021;49(1):69–75. https://doi.org/10.1186/s41182-021-00360-w
[36] Paul D, Sanap G, Shenoy S, Kalyane D, Kalia K, Tekade RK. Artificial intelligence in drug discovery and development. Drug Discov Today 2021;26(1):80–93. https://doi.org/10.1016/j.drudis.2020.10.010
[37] Naudé W. Artificial intelligence vs COVID-19: limitations, constraints and pitfalls. AI Soc 2020:1–5. https://doi.org/10.1007/s00146-020-00978-0
[38] Chen J, See KC. Artificial Intelligence for COVID-19: rapid review. J Med Internet Res 2020;22(10):e21476https://doi.org/10.2196/21476