AI-aided Precision Medicine against COVID-19: Strategic Areas of Research and Development

Enrico Santus, Nicola Marino, Davide Cirillo, Emmanuele Chersoni, Arnau Montagud, Antonella Santuccione Chada, Alfonso Valencia, Kevin Hughes, Charlotta Lindvall

Submitted to: Journal of Medical Internet Research
on: July 12, 2020

Table of Contents

Original Manuscript 4
Supplementary Files 31
Figures 32
Figure 1 33

AI-aided Precision Medicine against COVID-19: Strategic Areas of Research and Development

Enrico Santus[1]; Nicola Marino[2]; Davide Cirillo[3]; Emmanuele Chersoni[4]; Arnau Montagud[3]; Antonella Santuccione Chada[5]; Alfonso Valencia[3, 6]; Kevin Hughes[7]; Charlotta Lindvall[8]

[1]Division of Decision Science and Advanced Analytics Bayer Pharmaceuticals Whippany US
[2]Department of Medical and Surgical Sciences Università degli Studi di Foggia Foggia IT
[3]Barcelona Supercomputing Center Barcelona ES
[4]Department of Chinese and Bilingual Studies The Hong Kong Polytechnic University Hong Kong HK
[5]The Women's Brain Project Zurich CH
[6]ICREA Barcelona ES
[7]Massachusetts General Hospital Boston US
[8]Dana-Farber/Harvard Cancer Center Boston US

Corresponding Author:
Davide Cirillo
Barcelona Supercomputing Center
c/Jordi Girona, 29
Barcelona
ES

Abstract

Modern Artificial Intelligence (AI) technologies can play a key role in preventing and monitoring the present and the future epidemics. Here, we provide an overview of recently published literature on the COVID-19 pandemics in four strategic areas: (1) triage, diagnosis and risk prediction; (2) drug repurposing and development; (3) pharmacogenomics and vaccines; (4) mining of medical literature. In accordance to the newly proposed value-based care paradigm, we notice that AI-powered healthcare is essential to make public health systems able to efficiently handle future outbreaks and improve patient outcomes.

(JMIR Preprints 12/07/2020:22453)
DOI: https://doi.org/10.2196/preprints.22453

Preprint Settings

1) Would you like to publish your submitted manuscript as preprint?
Please make my preprint PDF available to anyone at any time (recommended).
Please make my preprint PDF available only to logged-in users; I understand that my title and abstract will remain visible to all users.
Only make the preprint title and abstract visible.
✓ **No, I do not wish to publish my submitted manuscript as a preprint.**

2) If accepted for publication in a JMIR journal, would you like the PDF to be visible to the public?
✓ **Yes, please make my accepted manuscript PDF available to anyone at any time (Recommended).**
Yes, but please make my accepted manuscript PDF available only to logged-in users; I understand that the title and abstract will remain v
Yes, but only make the title and abstract visible (see Important note, above). I understand that if I later pay to participate in <a href="http

Original Manuscript

Review

AI-aided Precision Medicine against COVID-19: Strategic Areas of Research and Development

Enrico Santus [1], Nicola Marino [2,3], Davide Cirillo [4], Emmanuele Chersoni [5], Arnau Montagud [4], Antonella Santuccione Chadha [3], Alfonso Valencia [4,6], Kevin Hughes [7] and Charlotta Lindvall [8]

1 Division of Decision Science and Advanced Analytics, Bayer Pharmaceuticals, 100 Bayer Boulevard, Whippany (NJ), United States

2 Department of Medical and Surgical Sciences, Università degli Studi di Foggia, via Antonio Gramsci 89, Foggia, Italy

3 Women's Brain Project, 8357 Guntershausen, Switzerland

4 Barcelona Supercomputing Center (BSC), C/ Jordi Girona, 29, 08034, Barcelona, Spain

5 Department of Chinese and Bilingual Studies, The Hong Kong Polytechnic University, 11 Yuk Choi Road, Hung Hom, Kowloon, Hong Kong, China

6 ICREA, Pg. Lluís Companys 23, 08010, Barcelona, Spain

7 Massachusetts General Hospital, 55 Fruit Street, Boston (MA), United States

8 Dana-Farber/Harvard Cancer Center, 450 Brookline Avenue, Boston (MA), United States

* corresponding author: davide.cirillo@bsc.es (Tel: +34 934137971)

ABSTRACT

Modern Artificial Intelligence (AI) technologies can play a key role in preventing and monitoring the present and the future epidemics. Here, we provide an overview of recently published literature on

the COVID-19 pandemics in four strategic areas: (1) triage, diagnosis and risk prediction; (2) drug repurposing and development; (3) pharmacogenomics and vaccines; (4) mining of medical literature. In accordance to the newly proposed value-based care paradigm, we notice that AI-powered healthcare is essential to make public health systems able to efficiently handle future outbreaks and improve patient outcomes.

INTRODUCTION

The recent pandemic of Severe Acute Respiratory Syndrome Coronavirus 2 (SARS-CoV-2) has highlighted the fragility of the healthcare systems to unexpected events, testing the endurance even of the top-performing ones [1]. As noted by several scholars, embracing artificial intelligence (AI) for healthcare optimization and outcome improvement is not an option anymore [2]. With respect to the pandemics, several studies have highlighted that the timely inclusion of AI-powered technologies would have accelerated the identification and the effective response to COVID-19 outbreaks. An example is the widely reported algorithm from a Canadian company, BlueDot, based on natural language processing and machine learning, which forecasted the emerging risk of a virus spread in Hubei province in late December 2019 by looking at news reports and airline ticketing [3].

Awareness of the benefit of employing AI to support and manage the COVID-19 crisis and its aftermath is increasing, in particular in the medical and research community. Notable examples of early AI-powered contributions include the discovery of relevant SARS-CoV-2 target proteins by Deep Mind’s AlphaFold [4] and the design by Infervision of a computer vision algorithm for coronavirus pneumonia detection in lung images [5].

Benefits do however come with technical challenges and related risks that still need to be properly assessed. For example, the absence of transparency and interpretability in AI models has hidden for a long time that the efficacy of these technologies is not equal across population groups. In particular, COVID-19 incidence and outcomes vary according to a large number of individual factors, including age, sex, ethnicity, health status, drug utilization, and others [6]. Sensitizing AI technologies to the diversity of the patient population and ensuring data security [7] is imperative to avoid biased decisions [8–10]. A crucial step to obtain robust, trustworthy and intelligible applications that account for demographic equity is therefore assessing potential biases in the resources used to train AI models for precision medicine [11].

As of today, AI systems are, regrettably, rarely endowed with robustness to class imbalances, such as sex and gender groups [12]. In this regard, sex differences in COVID-19 as well as gender-specific risk factors and socioeconomic burden have been recently highlighted in a case study by the European Commision [13]. Dataset multidimensionality that can fairly represent the population constitutes one of the main challenges for biobanking and cohort design efforts that collect different axes of health data [14]. In this regard, fair and broad data-collection systems are of primary importance. Two essential international references for COVID-19 genomic and medical data are the EMBL-EBI COVID-19 Data Portal [15] and the NIH National COVID Cohort Collaborative (N3C) [16]. The COVID-19 host genetics initiative [17] is an international collaborative undertaking to share resources for the investigation of the genetic determinants of COVID-19 susceptibility, severity, and outcomes [18]. The Coronavirus Pandemic Epidemiology (COPE) consortium aims to involve experts in the development of a personalized COVID-19 Symptom Tracker app that works as a real-time data-capture platform [6] which garnered over 2.8 million users in a few days. Moreover, COVID-19 sex-disaggregated data is collected by the Global Health 50/50 [19], an initiative housed at University College London advocating for gender equity.

Other ethical concerns include life/death decisions through risk prediction models, which may help to optimize resource allocation in time of scarcity. The application of non-optimal models may incur in the risk of worsening biases and exacerbating disparities for people with serious illnesses and different treatment priorities, also potentially causing the reduction in use of services rather than the achievement of the best patient care [20]. Still the power of prediction models is impressive and it may play an important role in the future, if properly exploited. For instance, a study from Cambridge University [21] shows how the use of secure AI operating on anonymised COVID-19 data can accurately predict the patient journey, allowing an optimal allocation of the resources and enabling well-informed and personalized healthcare decision-making. This is a particular important point, especially considering the difficulty in managing the increasing need for Intensive Care Units (ICU) during the COVID-19 pandemic peak [22,23].

The way in which the AI systems will be exploited is probably the most delicate topic in this adoption process, in particular if we refer to the decisional independence of the medical staff. As humans, in fact, clinicians are also affected by numerous cognitive biases, including the ‘confirmation bias’, which may lead them to give excessive importance to the evidence that supports the automated prediction (e.g. risk prediction, diagnosis, treatment suggestion) and ignore evidence that refutes it [8,24].

Despite the above mentioned concerns, there are numerous successful stories in the adoption of risk prediction models. For example, the Duke University adopted a system called Sepsis Watch that identifies in advance the inflammation leading to sepsis, one of the leading causes of hospital deaths. In two years from the tool introduction, the number of sepsis-induced patients has drastically decreased [25] thanks to three key elements: (1) the adaptation of the predictive model to a highly specific context; (2) scalability through integration with hospital workflows; and (3) the adopted user experience-based approach, which place clinicians and healthcare professionals at the center of the software development process, adhering with the Human in the Loop (HITL) paradigm [26,27].

The COVID-19 crisis is accelerating anticipated changes towards a stronger collaboration between computer science and medicine. In particular, the crisis has exposed the need for an increased scrutiny of the relationship between AI and patients as well as healthcare personnel under the lens of human and emotional needs, as demonstrated by the surge of mental health consequences of the pandemic [28] and the growing development of AI-based mental health apps and related digital tools [29]. Such aspects, together with others related to general data access and the use of AI for disease outcome prediction, are fueling the current debate about the convergence of AI and medicine [30,31] and the actionable realization of AI-powered innovations to fill the gap between technological research and medical practice, including applications in medical triage and advice, diagnostics and risk-adjusted paneling, population health management, digital devices integration, and many other areas [32]. Concerning this aspect, it is important to mention the recent publication of guidelines for the rigorous and transparent adoption of AI algorithm in the clinical practice: CONSORT-AI (Consolidated Standards of Reporting Trials – Artificial Intelligence) [33] and SPIRIT-AI (Standard Protocol Items: Recommendations for Interventional Trials – Artificial Intelligence) [34].

Translating patient data to successful therapies is the major ambition of implementing AI for health [35], especially in times of pandemic crisis, with the ultimate goal of achieving a successful bench-to-bedside model for better clinical decision-making [36,37]. In this work, we review some major examples of what AI has achieved during the global COVID-19 pandemic and the challenges that this technology and the medical community are currently facing in four main strategic areas of research and development (**Figure 1**): (1) triage, diagnosis and risk prediction; (2) drug repurposing and development; (3) pharmacogenomics and vaccines; (4) mining of medical literature.

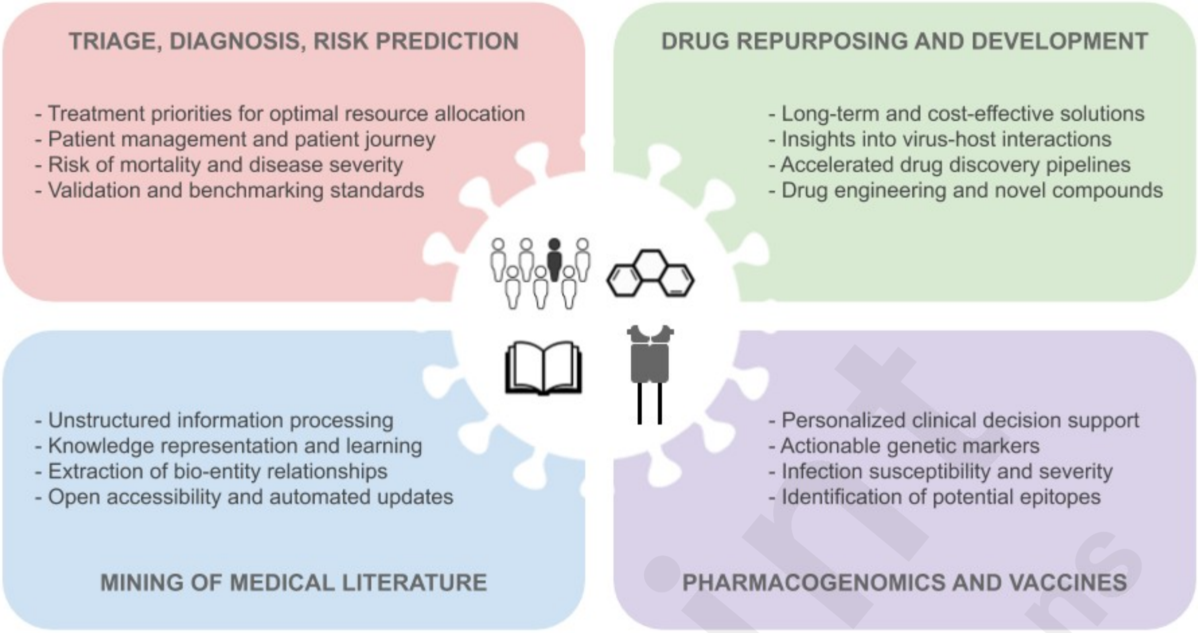

Figure 1. Main strategic areas of research and development for the realization of AI to fight COVID-19: triage, diagnosis and risk prediction; drug repurposing and development; pharmacogenomics and vaccines; mining of medical literature. The four panels report advantages and actionable solutions exhibited by the AI-aided precision medicine approaches surveyed in this work.

TRIAGE, DIAGNOSIS AND RISK PREDICTION

AI has been applied to determine treatment priorities in COVID-19 patients, or triage, and to better allocate limited resources. A group of researchers at the General Hospital of the People's Liberation Army (PLAGH), Beijing, China, has developed an online triage tool model (https://intensivecare.shinyapps.io/COVID19/) to manage suspected COVID-19 pneumonia in adult patients with fever [38]. By using clinical symptoms, routine laboratory tests and other clinical information available on admission as features, they trained a Logistic Regression with LASSO, obtaining an Area Under the ROC Curve (AUC) of 0.841 (sensitivity of 100% and specificity of 72.7%). Based on data from two hospitals in Wenzhou, Zhejiang, China, the authors of another recent study used an entropy-based feature selection approach: they modeled combinations of clinical features that could identify initial presentation patients at higher risk of developing severe illness, with an accuracy of 80% [39]. Results show that a mildly elevated alanine aminotransferase (ALT), the presence of myalgias (body aches), and an elevated hemoglobin level (red blood cells), in this order, are predictive of the later development of Acute Respiratory Distress Syndrome (ARDS).

A thorough study on risk prediction was carried out at the University of Cambridge based on the development of a proof of concept system to model the full patient journey through risk prediction models [21]. Identifying the risk of mortality and ICU/ventilator need, the system aimed at making doctors able to answer questions such as: (1) which patients are most likely to need ventilators within a week? (2) how many free ICU beds in the hospital are we likely to have in a week from now? and (3) which of two patients will get more benefits from going on a ventilator today? The predictive models showed accuracies between 77% for ventilator need to 83% for ICU admission and 87% for mortality.

Risk prediction models are not new to the AI-aided healthcare approach. They have been already successfully utilized for tasks such as predicting the risk of developing cancer [40,41] or identifying which patients are likely to benefit from heart-related procedures [42]. However, the COVID-19 crisis has accelerated their utilization. A recent work by Wynants and collaborators [43] has screened 14,217 titles about the pandemics from PubMed and Embase (Ovid, arXiv, medRxiv and bioRxiv), finding over 107 studies describing 145 prediction models. Among them, 4 models were aimed at identifying people at risk and 50 to predict the mortality risk, progression to severe disease, intensive care unit admission, ventilation, intubation or length of hospital stay. Not only do these models provide interesting results, but they also inform about the most valuable predictors: age, body temperature, lymphocyte count, lung imaging features and so on. Despite this, these models cannot be directly applied in the clinical setting without further validation, in order to guarantee data and experiment transparency and robustness, together with decision interpretability and model generalizability.

The remaining 91 models from this study were dedicated to the diagnosis of COVID-19, 60 of which exploited medical imaging. This number clearly shows that diagnosis is another important field for the application of AI techniques [44], with digital pathology exhibiting high effectiveness. In particular, Convolutional Neural Networks (CNNs) have been supporting radiologists in their expert decisions [45,46]. In a recent study, a CNN was trained to automatically learn patterns related to COVID-19 (i.e. ground-glass opacities, multifocal patchy consolidation, and/or interstitial changes with a predominantly peripheral distribution), achieving an AUC of 0.996 (with a sensitivity of 98.2% and a specificity of 92.2%) and outperforming the Reverse Transcription Polymerase Chain Reaction (RT-PCR), which also suffers from a significant time lag. In addition to accuracy, these approaches put the speed of the diagnosis on the table: CNNs can analyze up to 500 images in a few seconds, while radiologists would need hours.

Although chest computed tomography (CT) scans represent a commonly exploited source of information to train AI to rule out SARS-CoV-2 infection, the rapid detection of patients with COVID-19 can greatly benefit from learning approaches that utilize heterogeneous types of data. In this regard, it is crucial to consider the importance of training CNNs in a correct gender balance in medical imaging datasets to avoid producing distorted classifications for assisted diagnosis [12]. Moreover, it is crucial to rely on robust validation strategies to assess the generalization of the model to other datasets and populations, and high-quality benchmarking [47,48].

Indeed, AI can exploit multidimensional data, including the series of epidemiological, clinical, biological and radiological criteria defined by the WHO [49]. In a collaboration between researchers at hospitals in China and in the USA, CNN and other machine learning methods (support vector machine, random forest and multilayer perceptron classifiers) have been used to model and integrate CT scans and clinical information for diagnostic purposes [44]. The joint model that uses both sources of information achieved a 0.92 AUC (84.3% sensitivity, 82.8% specificity), outperforming the individual models. Moreover, the models allowed the identification of age, viral exposure, fever, cough and cough with sputum, and white blood cell counts, as the main features associated with SARS-CoV-2 status.

Recently, the NIH National Institute of Biomedical Imaging and Bioengineering (NIBIB) has launched the Medical Imaging and Data Resource Center (MIDRC) with the goal of coupling AI and medical imaging for COVID-19 early detection early detection and personalized therapies [50].

AI has also been utilized to identify patients at higher risk of mortality. Researchers at the Tongji Hospital, Wuhan, China, have screened Electronic Health Records (EHRs) of 375 discharged patients to use clinical measurements as features and have trained a gradient boosted decision tree model to predict the mortality risk [51]. The accuracy of the system was 93%. Its utilization would make it possible for physicians to immediately identify the critical cases and act accordingly. The model was also able to detect three key clinical features, that is, lactic dehydrogenase (LDH), lymphocyte, and High-sensitivity C-reactive protein (hs-CRP).

DRUG REPURPOSING AND DEVELOPMENT -

While triage, diagnosis and risk prediction are three of the most relevant tasks that AI has helped with during the peaks of the pandemic, other objectives are currently being addressed for long-term solutions. Among them are target selection for drug repurposing [52] and approaches for drug development, including de-novo drug design [53].

Drug repurposing consists of identifying existing drugs that could be effective in addressing proteins that are targeted by the virus. Recently, 332 high-confidence SARS-CoV-2 protein–human protein interactions have been experimentally identified, as well as 69 ligands, comprising FDA-approved drugs and compounds in preclinical and clinical trials, which specifically target these interactions [54]. Understanding which proteins and pathways the virus targets during the infection in the host is crucial for the development of AI systems for drug repurposing.

For instance, algorithms modeling the interaction between drugs and proteins have helped to identify baricitinib, which was previously used for the treatment of arthritis, as a useful drug against COVID-19 [55]. This drug does indeed inhibit the proteins that help the virus to penetrate the host cell. Thanks to approaches that exploit the computational identification of relations between existing drugs and target molecules, a research published by a team of Korean and American scientists has allowed the identification of FDA-approved antivirals that could potentially target the key proteins for COVID-19 [56].

The molecular processes of the virus-host interactions have been recently reconstructed in an international effort, coordinated by domain experts, called the COVID-19 Disease Map project [57]. The project aims to maintain an open-access resource for the continous curated integration of data and knowledge bases to support computational analysis and disease modeling. It represents a milestone of paramount importance for the development of AI systems for SARS-CoV-2 and their comparison with models of other coronaviruses. Moreover, by providing information about the intermolecular wiring of the virus-host interactions, the project enables network-based AI modeling for drug repurposing in COVID-19, which has recently shown promising results using network diffusion and network proximity [58]. Moreover, deep neural networks largely employed in Natural language processing (NLP), such as transformers architectures, have also been proposed for COVID-19 drug repurposing [59].

In the field of drug development, i.e. the pharmacotherapeutic course of a newly identified lead compound, computational models have been proven extremely successful in facilitating a quicker, cheaper and more effective development of new drugs [60]. For instance, AI can map multidimensional characteristics of proteins to greatly speed up the research process in comparison to traditional methodologies such as X-ray crystallography. In this regard, AI is crucial in optimizing drug discovery pipelines and improving outcomes of drug development, which is estimated to have a cost of US $2.6-billion [60].

Structural modeling and chemoinformatics methods for COVID-19 (e.g., docking-based binding-conformation studies of small molecules to target human or viral proteins) can greatly benefit from AI solutions. For instance, AI-based approaches have been used to infer structural similarities among molecules, such as algorithms able to model the graphical structure of chemical compounds through Graph Convolutional Networks (GCN) or other approaches [61]. AI systems can also leverage knowledge about protein sequences to infer the activity of similar ones. As previously mentioned, Google DeepMind has managed to predict the structure of six proteins that are targeted by SARS-Cov-2, namely the SARS-CoV-2 membrane protein, Nsp2, Nsp4, Nsp6, and Papain-like proteinase (C terminal domain) [62]. The deep learning approach uses amino acid features from similar sequences, based on multiple sequence alignment, to infer the distribution of structural distances to predict the protein structures [63].

Finally, AI can also be used to synthetically generate new molecules, such as new chemical compounds. For instance, the biotech company Insilico Medicine used reinforcement learning (RL) to model small molecules and identify those that inhibit specific infection pathways. The team created a generative chemistry pipeline to design novel SARS-CoV-2 inhibitors to later be synthesized and tested. The pipeline employs a large array of generative models, including autoencoders, generative adversarial networks (GANs), and genetic algorithms, which are optimized with RL [53].

PHARMACOGENOMICS AND VACCINES

Pharmacogenomics, that is the study of the role of genomic characteristics of an individual in drug response, represents a key gateway to personalized medicine [64–66]. Although the translation of the genomic information into clinical practice is recognized as one of the most challenging aspects of the future of medicine [67], the information about the genetic makeup of individual patients has the potential to guide clinical decision-support and to facilitate biomedical research in many different areas. For instance, genomics can inform drug discovery by providing simultaneous insights into the disease mechanisms and into potential targets for treating individual patients [68].

Pharmacogenomics approaches to COVID-19 are still in their infancy. Indeed, while the SARS-CoV2 genome was published in draft on January 10, 2020 [69], and real-time tracking of the pathogen evolution is now available (https://nextstrain.org/), much less genomic information is currently available about the host. A number of studies focus on genetic variations associated with infection susceptibility to infection and clinical manifestations, including human leukocyte antigen (HLA) variants in the UK Biobank population-based cohort [70] and ACE2 variants in the Italian population [71]. Retrospective and prospective studies focusing on COVID-19 disease susceptibility and severity have been collected by the COVID-19 Host Genetics Initiative [18].

Despite the absence of direct evidence of pharmacogenomics data in COVID-19 patients, the related literature for COVID-19 therapies, including hydroxychloroquine, ribavirin, baricitinib, and others, has been recently surveyed [72]. Potential actionable genetic markers have been reported, namely several genetic variants that are able to alter the pharmacokinetics of drugs that may affect the response to COVID-19 treatments. Importantly, as age, race, gender, and comorbidities have been associated with COVID-19 risk [73], these factors are deemed warranted to assess their role in the variation of treatment responses and need further investigation.

Population genetics is also needed to better understand the association between genetic variability and COVID-19. The importance and complexity of population genetic information, such as genome-wide association studies (GWAS), for drug discovery is exemplified by a study that shows that 8% of drugs approved by the US Food and Drug Administration (FDA) target molecules with genetic support, while only 2% of Phase 1 drugs are genetically supported [74]. Despite such low rates, GWAS studies can help to identify therapeutics that can be repurposed to treat individuals affected by diseases that are mechanistically related to those for which the drugs were developed [75]. Insights from GWAS can also inform about better patient management and therapy, such as the case of the variants in six genes on chromosome 3, namely SLC6A20, LZTFL1, CCR9, FYCO1, CXCR6, and XCR1, which have been recently associated with severe COVID-19 with respiratory failure [76].

Understanding population genetic heterogeneity is crucial for vaccine design, in particular as concerned the individual variability of the major histocompatibility complex (MHC-I and MHC-II) proteins, encoded by HLA gene, which present SARS-CoV-2 epitopes to the immune system. Such individual variability, coupled with the importance of cellular immunity in the severity of the response to the infection, makes the identification of actionable targets for COVID-19 vaccines a challenging endeavor. AI models for COVID-19 vaccine development focus on the prediction of potential epitopes using a variety of techniques, such as deep docking [77], Long Short-Term

Memory (LSTM) networks [78], eXtreme Gradient Boosting (XGBoost) [79], as well as approaches that account for different HLA alleles by combining several existing machine learning tools [80]. A recent survey of AI-based approaches to COVID-19 vaccine design [81] suggests that the most popular candidate is the SARS-CoV-2 spike protein (S-protein), which initiates the interaction with the host through the attachment to the ACE2 receptor [82].

MINING OF MEDICAL LITERATURE

The staggering rate of publications about COVID-19, both in the form of pre-prints and peer-reviewed articles, is posing unprecedented challenges to the knowledge acquisition and the information quality assessment process. A large part of contents is produced by humans for humans, in the form of free texts, where crucial pieces of information end up being buried. Because free text is not intelligible by machines, human intervention is required to identify the relevant pieces of information from the publications and turn it into a tabular form. Recent developments in Natural Language Processing (NLP) techniques have helped the automation of this process through machine learning and, in particular, deep learning algorithms [83,84]. Symptoms, patient demographics, clinical data, algorithms, performance, limitations are nowadays identifiable in the texts by properly trained models, which can obtain comparable accuracies to humans at a much faster rate, making it finally possible to monitor the enormous volume of produced literature [85]. The resulting structured data can be exploited to enrich knowledge graphs (KGs) [86–88], which provide a mean to represent and formalize structured information [86,89], analytical, relational and inferential investigations and fill the knowledge gaps in the community. Moreover, to rationalise the immense quantity of information on COVID-19, new algorithms can generate low-dimensional representations of the KGs, allowing researchers for clustering and classification [86,90]. We list here representative KG efforts that have been directed at the fight against COVID-19 (see **Table 1**).

Project name	**URL**
KG-Covid-19 Knowledge Graph Hub	https://github.com/Knowledge-Graph-Hub/kg-covid-19
COVID-19 Community Project	https://github.com/covid-19-net/covid-19-community
COVID-KG	https://github.com/GillesVandewiele/COVID-KG/

CovidGraph	https://covidgraph.org/
COVID-19 Miner	https://github.com/rupertoverall/covid19-miner
COVID-19 Knowledge Graph	https://bikmi.covid19-knowledgespace.de/
Taxila-COVID-19	http://covid19.taxila.io/

Table 1. Knowledge graph (KG) resources for COVID-19.

The KG-Covid-19 Knowledge Graph Hub project is the first Knowledge Graph Hub (KG Hub) dedicated to COVID-19. A KGHub is a software to download and transform data to a central location for building KGs from different combinations of data sources. The Covid-19 KGHub downloads and transforms data from more than 50 different COVID-19 databases of drugs, genes, proteins, ontologies, diseases, phenotypes and publications, and emits a KG that can be used for machine learning.

The COVID-19 Community Project is a community-based KG that links heterogeneous datasets about COVID-19, in three main areas: the host, the virus and the cellular environment. These KGs use several publicly-available datasets, such as the CORD-19 dataset, a set of over 51,000 scholarly articles about coronaviruses [91].

Other notable databases used in KGs are the COVID-19 Data Portal (see Introduction) and the COVID-19 crowd-generated gene and drug set library (https://amp.pharm.mssm.edu/covid19/). One of the tools that uses these is the COVID-KG, which graph embeddings of the entities in the KG, such as papers, authors or journals [92].

CovidGraph is a collaboration of researchers to build a research and communication platform that encompasses over 40,000 publications, case statistics, genes and functions, molecular data and much more. The output is a KG in which entity relationships can be found and new pieces of literature discovered. Another tool that uses the CORD-19 dataset is COVID-19 Miner, which provides access to a database of interactions between genes/proteins, chemicals and biological processes related to the SARS-CoV-2 virus automatically extracted using NLP from the CORD-19 dataset and daily-updated manuscripts from medRxiv and bioRxiv [93].

Furthermore, COVID-19 Knowledge Graph [86] is an intent to lay the foundation for a comprehensive and interactive KG in the context of COVID-19 that connects causes and effects and enables users to completely explore the contained information. Data comes from papers from PubMed and preprints from bioRxiv, chemRxiv, medRxiv, PrePrints and Researchsquare. Lastly, Taxila-COVID-19 is an AI and NLP system which uses thousands of COVID-19-related publications, clinical trials and other relevant sources to enable users to search and analyze COVID-19 literature. Publications and data are automatically updated.

DISCUSSION

The COVID-19 pandemic has caused some of the most significant challenges that national healthcare systems, even well-developed ones, have had to face in recent human history. Components of such systems include not only hospital care facilities but also practitioners and family doctors, retirement and nursing homes, families and communities, which have all been affected in an unprecedented way. This has impacted the society as a whole, with lockdown policies taken to reduce the hospital overloads, and has also had social and economic consequences, which have been more severe for minorities and vulnerable groups [94]. Moreover, this pandemic is a worldwide phenomenon in the age of the social media and of the Web 2.0, with the consequence of facilitating the spreading of misinformation and fake news from multiple actors, and with no way for the average Internet user to check the reliability of the sources. Nevertheless, the lessons we have learnt—and some we are still learning—from the COVID-19 crisis are shaping the future of our community towards the role of technology in facilitating a better understanding of the complex issues surrounding the disease and all the factors that influence its impact and significance for public health.

In this work we surveyed several examples of AI stepping up to advance research and medical practice and prepare governments for responsible actions in future similar scenarios. AI has been proven to improve efficiency in the healthcare pipeline and to efficiently deliver models for outbreak analytics and detection, prevention and early intervention, and decision-making. In particular, we showed the role that such innovative technologies play as a lifeline for present and future epidemics, and the unparalleled opportunity for our community to utilize AI advancement to fill the gap between translational research and medical practice. Finally, in addition to the immediate medical applications of the AI implementations that we survey, it is worth mentioning the importance of Natural Language Processing for monitoring the quality of the information available to the public and fighting fake news [95–97].

Thanks to the availability of large amounts of data (i.e. Big data) and high-performance computing (HPC), the fight against the novel coronavirus can leverage the support of artificial intelligence (AI), as demonstrated by initiatives such as the COVID-19 HPC Consortium (https://covid19-hpc-consortium.org/). This technology allows us to address, at a much higher speed and a comparable performance, complex tasks that cannot be executed by humans, who can now focus on more intelligence-demanding activities, such as emotional intelligence and human-to-human bonding [98].

While AI is traditionally trained on large datasets for identifying population-level patterns (i.e., common characteristics among people belonging to some clinical classes), recent efforts have promoted the utilization of this technology in conjunction to the principles of precision medicine, to substitute the “average patient” [41] with a real individual, based on geographical and socioeconomic signature as well as genetic, epigenetic and other molecular profiles [99]. Under this paradigm, AI is meant to empower clinicians to tailor interventions [100] (whether preventive or therapeutic) to the nuanced—and often unique—features of every human being [101]. To this end, multidimensional datasets, such as the variety of data modalities that are currently collected and modelled for COVID-19 [102–104], capture individual genetic, biochemical, physiological, environmental, and behavioral variations [105] that may interfere with the development, progression and treatment of a disease. Thanks to the drop in price of sequencing the human genome (from billions to hundreds of dollars in 30 years [106]), it is now possible to exploit AI to study the phenotypic, genotypic and environmental correlations with diseases [107]. With such an analysis, not only AI can predict the risk of a subject to develop a disease, but it can also estimate the likelihood of success for a treatment. In the case of COVID-19, this could lead to a better allocation of the resources and to an improved match between treatments and patients, with a consequent improved outcome for preventive and therapeutic interventions. AI-aided precision medicine therefore connects some of the key benefits for a sustainable and effective healthcare system: efficiency, efficacy and safety assessment [30].

AI is recognized as a necessity to achieve precision medicine in COVID-19. The current crisis has highlighted that there is still a huge amount of work needed to exploit AI-based solutions to their full potential in order to transform healthcare. In particular, AI implementation in the clinical setting is still far to be complete [107]. The highly fragmented and diverse healthcare systems, the absence of a protocol in documenting patient data, the ethical constraints (such as privacy) and the limitations of AI itself (e.g. bias and non-interpretability) still represent serious challenges to extensive AI adoption [108].

Acknowledgments

This work is written on behalf of the Women’s Brain Project (WBP) (www.womensbrainproject.com/), an international organization advocating for women’s brain and mental health through scientific research, debate and public engagement. The corresponding author had final responsibility for the decision to submit for publication.

Competing interest statement: No conflict of interest.

Authors’ contributions

ES and ASC conceived the study. NM, DC, AM and EC assisted with content selection, design, and revision. AV, KH and CL supervised the project. All the authors contributed writing and editing the manuscript.

References

1. Legido-Quigley H, Asgari N, Teo YY, Leung GM, Oshitani H, Fukuda K, Cook AR, Hsu LY, Shibuya K, Heymann D. Are high-performing health systems resilient against the COVID-19 epidemic? Lancet [Internet] 2020 Mar;395(10227):848–850. [doi: 10.1016/S0140-6736(20)30551-1]

2. McCall B. COVID-19 and artificial intelligence: protecting health-care workers and curbing the spread. The Lancet Digital Health [Internet] 2020 Apr;2(4):e166–e167. [doi: 10.1016/S2589-7500(20)30054-6]

3. Niiler E. An AI Epidemiologist Sent the First Alerts of the Coronavirus [Internet]. Wired. WIRED; 2020 [cited 2020 Jul 7]. Available from: https://www.wired.com/story/ai-epidemiologist-wuhan-public-health-warnings/

4. Computational predictions of protein structures associated with COVID-19 [Internet]. Deepmind. [cited 2020 Jul 7]. Available from: https://deepmind.com/research/open-source/computational-predictions-of-protein-structures-

associated-with-COVID-19

5. Lan L, Xu D, Ye G, Xia C, Wang S, Li Y, Xu H. Positive RT-PCR Test Results in Patients Recovered From COVID-19. JAMA [Internet] 2020 Feb 27; PMID:32105304

6. Drew DA, Nguyen LH, Steves CJ, Menni C, Freydin M, Varsavsky T, Sudre CH, Cardoso MJ, Ourselin S, Wolf J, Spector TD, Chan AT, COPE Consortium. Rapid implementation of mobile technology for real-time epidemiology of COVID-19. Science [Internet] 2020 May 5; PMID:32371477

7. Shaban-Nejad A, Michalowski M, Buckeridge DL. Health intelligence: how artificial intelligence transforms population and personalized health. NPJ Digit Med [Internet] 2018 Oct 2;1:53. PMID:31304332

8. Challen R, Denny J, Pitt M, Gompels L, Edwards T, Tsaneva-Atanasova K. Artificial intelligence, bias and clinical safety. BMJ Qual Saf [Internet] 2019 Mar;28(3):231–237. PMID:30636200

9. Parikh RB, Teeple S, Navathe AS. Addressing Bias in Artificial Intelligence in Health Care. JAMA [Internet] 2019 Nov 22; PMID:31755905

10. Nelson GS. Bias in Artificial Intelligence. N C Med J [Internet] 2019 Jul;80(4):220–222. PMID:31278182

11. Cirillo D, Catuara-Solarz S, Morey C, Guney E, Subirats L, Mellino S, Gigante A, Valencia A, Rementeria MJ, Chadha AS, Mavridis N. Sex and gender differences and biases in artificial intelligence for biomedicine and healthcare. NPJ Digit Med [Internet] 2020 Jun 1;3:81. PMID:32529043

12. Larrazabal AJ, Nieto N, Peterson V, Milone DH, Ferrante E. Gender imbalance in medical imaging datasets produces biased classifiers for computer-aided diagnosis. Proc Natl Acad Sci U S A [Internet] 2020 May 26; PMID:32457147

13. The impact of sex and gender in the COVID-19 pandemic [Internet]. European Commission - European Commission. 2020 [cited 2020 Jul 7]. Available from: https://ec.europa.eu/info/publications/impact-sex-and-gender-covid-19-pandemic_en

14. Shilo S, Rossman H, Segal E. Axes of a revolution: challenges and promises of big data in healthcare. Nat Med [Internet] 2020 Jan;26(1):29–38. PMID:31932803

15. COVID-19 Data Portal [Internet]. [cited 2020 Oct 6]. Available from: https://www.covid19dataportal.org/

16. National COVID Cohort Collaborative (N3C) [Internet]. 2020 [cited 2020 Oct 6]. Available from: https://ncats.nih.gov/n3c

17. Covid19 Host Genetics Initiative [Internet]. [cited 2020 Oct 6]. Available from:

https://www.covid19hg.org/

18. COVID-19 Host Genetics Initiative. The COVID-19 Host Genetics Initiative, a global initiative to elucidate the role of host genetic factors in susceptibility and severity of the SARS-CoV-2 virus pandemic. Eur J Hum Genet [Internet] 2020 May 13; PMID:32404885

19. The Sex, Gender and COVID-19 Project [Internet]. [cited 2020 Oct 6]. Available from: https://globalhealth5050.org/the-sex-gender-and-covid-19-project/

20. Lindvall C, Cassel CK, Pantilat SZ, DeCamp M. Ethical Considerations In The Use Of AI Mortality Predictions In The Care Of People With Serious Illness [Internet]. [cited 2020 Oct 4]. Available from: https://www.healthaffairs.org/do/10.1377/hblog20200911.401376/full/

21. Progress using COVID-19 patient data to train machine learning models for healthcare [Internet]. University of Cambridge. 2020 [cited 2020 Jul 7]. Available from: https://www.cam.ac.uk/research/news/progress-using-covid-19-patient-data-to-train-machine-learning-models-for-healthcare

22. Rahmatizadeh S, Valizadeh-Haghi S, Dabbagh A. The role of Artificial Intelligence in Management of Critical COVID-19 patients. Journal of Cellular & Molecular Anesthesia [Internet] 2020 Apr 9 [cited 2020 Oct 6];5(1):16–22. [doi: 10.22037/jcma.v5i1.29752]

23. Grasselli G, Pesenti A, Cecconi M. Critical Care Utilization for the COVID-19 Outbreak in Lombardy, Italy: Early Experience and Forecast During an Emergency Response. JAMA [Internet] American Medical Association; 2020 Apr 28 [cited 2020 Oct 6];323(16):1545–1546. [doi: 10.1001/jama.2020.4031]

24. Dawson NV, Arkes HR. Systematic errors in medical decision making: judgment limitations. J Gen Intern Med [Internet] 1987 May;2(3):183–187. PMID:3295150

25. Sendak MP, Ratliff W, Sarro D, Alderton E, Futoma J, Gao M, Nichols M, Revoir M, Yashar F, Miller C, Kester K, Sandhu S, Corey K, Brajer N, Tan C, Lin A, Brown T, Engelbosch S, Anstrom K, Elish MC, Heller K, Donohoe R, Theiling J, Poon E, Balu S, Bedoya A, O'Brien C. Real-World Integration of a Sepsis Deep Learning Technology Into Routine Clinical Care: Implementation Study. JMIR Med Inform [Internet] 2020 Jul 15;8(7):e15182. PMID:32673244

26. Verghese A, Shah NH, Harrington RA. What This Computer Needs Is a Physician: Humanism and Artificial Intelligence. JAMA [Internet] 2018 Jan 2;319(1):19–20. PMID:29261830

27. Patel BN, Rosenberg L, Willcox G, Baltaxe D, Lyons M, Irvin J, Rajpurkar P, Amrhein T, Gupta R, Halabi S, Langlotz C, Lo E, Mammarappallil J, Mariano AJ, Riley G, Seekins J, Shen L, Zucker E, Lungren M. Human-machine partnership with artificial intelligence for chest radiograph diagnosis. NPJ Digit Med [Internet] 2019 Nov 18;2:111. PMID:31754637

28. Vindegaard N, Benros ME. COVID-19 pandemic and mental health consequences: Systematic review of the current evidence. Brain Behav Immun [Internet] 2020 May 30; PMID:32485289

29. Figueroa CA, Aguilera A. The Need for a Mental Health Technology Revolution in the COVID-19 Pandemic. Front Psychiatry [Internet] 2020 Jun 3;11:523. PMID:32581891

30. Topol EJ. High-performance medicine: the convergence of human and artificial intelligence. Nat Med [Internet] 2019 Jan;25(1):44–56. PMID:30617339

31. Panch T, Mattie H, Celi LA. The “inconvenient truth” about AI in healthcare. NPJ Digit Med [Internet] 2019 Aug 16;2:77. PMID:31453372

32. Lin SY, Mahoney MR, Sinsky CA. Ten Ways Artificial Intelligence Will Transform Primary Care. J Gen Intern Med [Internet] 2019 Aug;34(8):1626–1630. PMID:31090027

33. Liu X, Cruz Rivera S, Moher D, Calvert MJ, Denniston AK, SPIRIT-AI and CONSORT-AI Working Group. Reporting guidelines for clinical trial reports for interventions involving artificial intelligence: the CONSORT-AI extension. Nat Med [Internet] 2020 Sep;26(9):1364–1374. PMID:32908283

34. Rivera SC, Liu X, Chan A-W, Denniston AK, Calvert MJ. Guidelines for clinical trial protocols for interventions involving artificial intelligence: the SPIRIT-AI extension. Nat Med [Internet] Nature Publishing Group; 2020 Sep 9 [cited 2020 Oct 6];26(9):1351–1363. [doi: 10.1038/s41591-020-1037-7]

35. Filipp FV. Opportunities for Artificial Intelligence in Advancing Precision Medicine. Curr Genet Med Rep [Internet] 2019 Dec;7(4):208–213. PMID:31871830

36. Zeggini E, Gloyn AL, Barton AC, Wain LV. Translational genomics and precision medicine: Moving from the lab to the clinic. Science [Internet] 2019 Sep 27;365(6460):1409–1413. PMID:31604268

37. Gruson D, Helleputte T, Rousseau P, Gruson D. Data science, artificial intelligence, and machine learning: Opportunities for laboratory medicine and the value of positive regulation. Clin Biochem [Internet] 2019 Jul;69:1–7. PMID:31022391

38. Feng C, Huang Z, Wang L, Chen X, Zhai Y, Zhu F, Chen H, Wang Y, Su X, Huang S, Tian L, Zhu W, Sun W, Zhang L, Han Q, Zhang J, Pan F, Chen L, Zhu Z, Xiao H, Liu Y, Liu G, Chen W, Li T. A Novel Triage Tool of Artificial Intelligence Assisted Diagnosis Aid System for Suspected COVID-19 pneumonia In Fever Clinics. Emergency Medicine. medRxiv; 2020.

39. Jiang X, Coffee M, Bari A, Wang J, Jiang X, Huang J, Shi J, Dai J, Cai J, Zhang T, Wu Z, He G, Huang Y. Towards an Artificial Intelligence Framework for Data-Driven Prediction of Coronavirus Clinical Severity. Computers, Materials & Continua [Internet] 2020;62(3):537–551. [doi: 10.32604/cmc.2020.010691]

40. Freedman AN, Seminara D, Gail MH, Hartge P, Colditz GA, Ballard-Barbash R, Pfeiffer RM. Cancer risk prediction models: a workshop on development, evaluation, and application. J Natl Cancer Inst [Internet] 2005 May 18;97(10):715–723. PMID:15900041

41. Yala A, Lehman C, Schuster T, Portnoi T, Barzilay R. A Deep Learning Mammography-based Model for Improved Breast Cancer Risk Prediction. Radiology [Internet] 2019 Jul;292(1):60–66. PMID:31063083

42. Hu S-Y, Santus E, Forsyth AW, Malhotra D, Haimson J, Chatterjee NA, Kramer DB, Barzilay R, Tulsky JA, Lindvall C. Can machine learning improve patient selection for cardiac resynchronization therapy? PLoS One [Internet] 2019 Oct 3;14(10):e0222397. PMID:31581234

43. Wynants L, Van Calster B, Collins GS, Riley RD, Heinze G, Schuit E, Bonten MMJ, Damen JAA, Debray TPA, De Vos M, Dhiman P, Haller MC, Harhay MO, Henckaerts L, Kreuzberger N, Lohman A, Luijken K, Ma J, Andaur CL, Reitsma JB, Sergeant JC, Shi C, Skoetz N, Smits LJM, Snell KIE, Sperrin M, Spijker R, Steyerberg EW, Takada T, van Kuijk SMJ, van Royen FS, Wallisch C, Hooft L, Moons KGM, van Smeden M. Prediction models for diagnosis and prognosis of covid-19 infection: systematic review and critical appraisal. BMJ [Internet] 2020 Apr 7;369:m1328. PMID:32265220

44. Mei X, Lee H-C, Diao K-Y, Huang M, Lin B, Liu C, Xie Z, Ma Y, Robson PM, Chung M, Bernheim A, Mani V, Calcagno C, Li K, Li S, Shan H, Lv J, Zhao T, Xia J, Long Q, Steinberger S, Jacobi A, Deyer T, Luksza M, Liu F, Little BP, Fayad ZA, Yang Y. Artificial intelligence-enabled rapid diagnosis of patients with COVID-19. Nat Med [Internet] 2020 May 19; PMID:32427924

45. Butt C, Gill J, Chun D, Babu BA. Deep learning system to screen coronavirus disease 2019 pneumonia. Appl Intell [Internet] 2020 Apr 22;9:291. [doi: 10.1007/s10489-020-01714-3]

46. Jin C, Chen W, Cao Y, Xu Z, Zhang X, Deng L, Zheng C, Zhou J, Shi H, Feng J. Development and Evaluation of an AI System for COVID-19 Diagnosis. Radiology and Imaging. medRxiv; 2020.

47. Schaffter T, Buist DSM, Lee CI, Nikulin Y, Ribli D, Guan Y, Lotter W, Jie Z, Du H, Wang S, Feng J, Feng M, Kim H-E, Albiol F, Albiol A, Morrell S, Wojna Z, Ahsen ME, Asif U, Jimeno Yepes A, Yohanandan S, Rabinovici-Cohen S, Yi D, Hoff B, Yu T, Chaibub Neto E, Rubin DL, Lindholm P, Margolies LR, McBride RB, Rothstein JH, Sieh W, Ben-Ari R, Harrer S, Trister A, Friend S, Norman T, Sahiner B, Strand F, Guinney J, Stolovitzky G, and the DM DREAM Consortium, Mackey L, Cahoon J, Shen L, Sohn JH, Trivedi H, Shen Y, Buturovic L, Pereira JC, Cardoso JS, Castro E, Kalleberg KT, Pelka O, Nedjar I, Geras KJ, Nensa F, Goan E, Koitka S, Caballero L, Cox DD, Krishnaswamy P, Pandey G, Friedrich CM, Perrin D, Fookes C, Shi B, Cardoso Negrie G, Kawczynski M, Cho K, Khoo CS, Lo JY, Sorensen AG, Jung H. Evaluation of Combined Artificial Intelligence and Radiologist Assessment to Interpret Screening Mammograms. JAMA Netw Open [Internet] 2020 Mar 2;3(3):e200265. PMID:32119094

48. Xiao L-S, Li P, Sun F, Zhang Y, Xu C, Zhu H, Cai F-Q, He Y-L, Zhang W-F, Ma S-C, Hu C, Gong M, Liu L, Shi W, Zhu H. Development and Validation of a Deep Learning-Based Model Using Computed Tomography Imaging for Predicting Disease Severity of Coronavirus Disease 2019. Front Bioeng Biotechnol [Internet] 2020 Jul 31;8:898. PMID:32850746

49. World Health Organization. Clinical management of severe acute respiratory infection (SARI) when COVID-19 disease is suspected. 2020; Available from: https://www.who.int/docs/default-source/coronaviruse/clinical-management-of-novel-cov.pdf?sfvrsn=bc7da517_2

50. NIH harnesses AI for COVID-19 diagnosis, treatment, and monitoring [Internet]. 2020 [cited 2020 Oct 6]. Available from: https://www.nih.gov/news-events/news-releases/nih-harnesses-ai-covid-19-diagnosis-treatment-monitoring

51. Yan L, Zhang H-T, Goncalves J, Xiao Y, Wang M, Guo Y, Sun C, Tang X, Jin L, Zhang M, Huang X, Xiao Y, Cao H, Chen Y, Ren T, Wang F, Xiao Y, Huang S, Tan X, Huang N, Jiao B, Zhang Y, Luo A, Mombaerts L, Jin J, Cao Z, Li S, Xu H, Yuan Y. A machine learning-based model for survival prediction in patients with severe COVID-19 infection. Epidemiology. medRxiv; 2020.

52. Senanayake SL. Drug repurposing strategies for COVID-19 [Internet]. Future Drug Discovery. 2020. [doi: 10.4155/fdd-2020-0010]

53. Zhavoronkov A, Aladinskiy V, Zhebrak A, Zagribelnyy B, Terentiev V, Bezrukov DS, Polykovskiy D, Shayakhmetov R, Filimonov A, Orekhov P, Yan Y, Popova O, Vanhaelen Q, Aliper A, Ivanenkov Y. Potential COVID-2019 3C-like Protease Inhibitors Designed Using Generative Deep Learning Approaches. 2020.

54. Gordon DE, Jang GM, Bouhaddou M, Xu J, Obernier K, White KM, O'Meara MJ, Rezelj VV, Guo JZ, Swaney DL, Tummino TA, Hüttenhain R, Kaake RM, Richards AL, Tutuncuoglu B, Foussard H, Batra J, Haas K, Modak M, Kim M, Haas P, Polacco BJ, Braberg H, Fabius JM, Eckhardt M, Soucheray M, Bennett MJ, Cakir M, McGregor MJ, Li Q, Meyer B, Roesch F, Vallet T, Mac Kain A, Miorin L, Moreno E, Naing ZZC, Zhou Y, Peng S, Shi Y, Zhang Z, Shen W, Kirby IT, Melnyk JE, Chorba JS, Lou K, Dai SA, Barrio-Hernandez I, Memon D, Hernandez-Armenta C, Lyu J, Mathy CJP, Perica T, Pilla KB, Ganesan SJ, Saltzberg DJ, Rakesh R, Liu X, Rosenthal SB, Calviello L, Venkataramanan S, Liboy-Lugo J, Lin Y, Huang X-P, Liu Y, Wankowicz SA, Bohn M, Safari M, Ugur FS, Koh C, Savar NS, Tran QD, Shengjuler D, Fletcher SJ, O'Neal MC, Cai Y, Chang JCJ, Broadhurst DJ, Klippsten S, Sharp PP, Wenzell NA, Kuzuoglu-Ozturk D, Wang H-Y, Trenker R, Young JM, Cavero DA, Hiatt J, Roth TL, Rathore U, Subramanian A, Noack J, Hubert M, Stroud RM, Frankel AD, Rosenberg OS, Verba KA, Agard DA, Ott M, Emerman M, Jura N, von Zastrow M, Verdin E, Ashworth A, Schwartz O, d'Enfert C, Mukherjee S, Jacobson M, Malik HS, Fujimori DG, Ideker T, Craik CS, Floor SN, Fraser JS, Gross JD, Sali A, Roth BL, Ruggero D, Taunton J, Kortemme T, Beltrao P, Vignuzzi M, García-Sastre A, Shokat KM, Shoichet BK, Krogan NJ. A SARS-CoV-2 protein interaction map reveals targets for drug repurposing. Nature [Internet] 2020 Jul;583(7816):459–468. PMID:32353859

55. Richardson P, Griffin I, Tucker C, Smith D, Oechsle O, Phelan A, Stebbing J. Baricitinib as potential treatment for 2019-nCoV acute respiratory disease. Lancet [Internet] 2020 Feb;395(10223):e30–e31. [doi: 10.1016/S0140-6736(20)30304-4]

56. Beck BR, Shin B, Choi Y, Park S, Kang K. Predicting commercially available antiviral drugs

that may act on the novel coronavirus (2019-nCoV), Wuhan, China through a drug-target interaction deep learning model. Microbiology. bioRxiv; 2020. p. 230.

57. Ostaszewski M, Mazein A, Gillespie ME, Kuperstein I, Niarakis A, Hermjakob H, Pico AR, Willighagen EL, Evelo CT, Hasenauer J, Schreiber F, Dräger A, Demir E, Wolkenhauer O, Furlong LI, Barillot E, Dopazo J, Orta-Resendiz A, Messina F, Valencia A, Funahashi A, Kitano H, Auffray C, Balling R, Schneider R. COVID-19 Disease Map, building a computational repository of SARS-CoV-2 virus-host interaction mechanisms. Sci Data [Internet] 2020 May 5;7(1):136. PMID:32371892

58. Gysi DM, Valle ÍD, Zitnik M, Ameli A, Gan X, Varol O, Sanchez H, Baron RM, Ghiassian D, Loscalzo J, Barabási A-L. Network Medicine Framework for Identifying Drug Repurposing Opportunities for COVID-19 [Internet]. 2020 [cited 2020 Jul 7]. Available from: http://arxiv.org/abs/2004.07229

59. 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 [Internet] 2020 Mar 30;18:784–790. PMID:32280433

60. Fleming N. How artificial intelligence is changing drug discovery. Nature [Internet] 2018 May;557(7707):S55–S57. PMID:29849160

61. Stokes JM, Yang K, Swanson K, Jin W, Cubillos-Ruiz A, Donghia NM, MacNair CR, French S, Carfrae LA, Bloom-Ackermann Z, Tran VM, Chiappino-Pepe A, Badran AH, Andrews IW, Chory EJ, Church GM, Brown ED, Jaakkola TS, Barzilay R, Collins JJ. A Deep Learning Approach to Antibiotic Discovery. Cell [Internet] 2020 Feb 20;180(4):688–702.e13. PMID:32084340

62. Computational predictions of protein structures associated with COVID-19 [Internet]. Deepmind. [cited 2020 Jul 7]. Available from: https://deepmind.com/research/open-source/computational-predictions-of-protein-structures-associated-with-COVID-19

63. Senior AW, Evans R, Jumper J, Kirkpatrick J, Sifre L, Green T, Qin C, Žídek A, Nelson AWR, Bridgland A, Penedones H, Petersen S, Simonyan K, Crossan S, Kohli P, Jones DT, Silver D, Kavukcuoglu K, Hassabis D. Improved protein structure prediction using potentials from deep learning. Nature [Internet] 2020 Jan;577(7792):706–710. PMID:31942072

64. Nelson MR, Johnson T, Warren L, Hughes AR, Chissoe SL, Xu C-F, Waterworth DM. The genetics of drug efficacy: opportunities and challenges. Nat Rev Genet [Internet] 2016 Apr;17(4):197–206. PMID:26972588

65. Tan-Koi WC, Leow PC, Teo YY. Applications of pharmacogenomics in regulatory science: a product life cycle review. Pharmacogenomics J [Internet] 2018 May 22;18(3):359–366. PMID:29205206

66. Weinshilboum RM, Wang L. Pharmacogenomics: Precision Medicine and Drug Response. Mayo Clin Proc [Internet] 2017 Nov;92(11):1711–1722. PMID:29101939

67. Lauschke VM, Ingelman-Sundberg M. Emerging strategies to bridge the gap between pharmacogenomic research and its clinical implementation. NPJ Genom Med [Internet] 2020 Mar 5;5:9. PMID:32194983

68. Silverman EK. Genetics of COPD. Annu Rev Physiol [Internet] 2020 Feb 10;82:413–431. PMID:31730394

69. Severe acute respiratory syndrome coronavirus 2 isolate Wuhan-Hu-1, co - Nucleotide - NCBI [Internet]. [cited 2020 Oct 4]. Available from: https://www.ncbi.nlm.nih.gov/nuccore/MN908947

70. Kachuri L, Francis SS, Morrison M, Bossé Y, Cavazos TB, Rashkin SR, Ziv E, Witte JS. The landscape of host genetic factors involved in infection to common viruses and SARS-CoV-2. Genetic and Genomic Medicine. medRxiv; 2020.

71. Benetti E, Tita R, Spiga O, Ciolfi A, Birolo G, Bruselles A, Doddato G, Giliberti A, Marconi C, Musacchia F, Pippucci T, Torella A, Trezza A, Valentino F, Baldassarri M, Brusco A, Asselta R, Mirella B, Furini S, Seri M, Nigro V, Matullo G, Tartaglia M, Mari F, Renieri A, Pinto A. ACE2 gene variants may underlie interindividual variability and susceptibility to COVID-19 in the Italian population. Genetic and Genomic Medicine. medRxiv; 2020.

72. Takahashi T, Luzum JA, Nicol MR, Jacobson PA. Pharmacogenomics of COVID-19 therapies. NPJ Genom Med [Internet] 2020 Aug 18;5:35. PMID:32864162

73. Wenham C, Smith J, Morgan R. COVID-19: the gendered impacts of the outbreak. Lancet [Internet] 2020 Mar;395(10227):846–848. [doi: 10.1016/S0140-6736(20)30526-2]

74. Nelson MR, Tipney H, Painter JL, Shen J, Nicoletti P, Shen Y, Floratos A, Sham PC, Li MJ, Wang J, Cardon LR, Whittaker JC, Sanseau P. The support of human genetic evidence for approved drug indications. Nat Genet [Internet] 2015 Aug;47(8):856–860. PMID:26121088

75. Dugger SA, Platt A, Goldstein DB. Drug development in the era of precision medicine. Nat Rev Drug Discov [Internet] 2018 Mar;17(3):183–196. PMID:29217837

76. Ellinghaus D, Degenhardt F, Bujanda L, Buti M, Albillos A, Invernizzi P, Fernández J, Prati D, Baselli G, Asselta R, Grimsrud MM, Milani C, Aziz F, Kässens J, May S, Wendorff M, Wienbrandt L, Uellendahl-Werth F, Zheng T, Yi X, de Pablo R, Chercoles AG, Palom A, Garcia-Fernandez A-E, Rodriguez-Frias F, Zanella A, Bandera A, Protti A, Aghemo A, Lleo A, Biondi A, Caballero-Garralda A, Gori A, Tanck A, Carreras Nolla A, Latiano A, Fracanzani AL, Peschuck A, Julià A, Pesenti A, Voza A, Jiménez D, Mateos B, Nafria Jimenez B, Quereda C, Paccapelo C, Gassner C, Angelini C, Cea C, Solier A, Pestaña D, Muñiz-Diaz E, Sandoval E, Paraboschi EM, Navas E, García Sánchez F, Ceriotti F, Martinelli-Boneschi F, Peyvandi F, Blasi F, Téllez L, Blanco-Grau A, Hemmrich-Stanisak G, Grasselli G, Costantino G, Cardamone G, Foti G, Aneli S, Kurihara H, ElAbd H, My I, Galván-Femenia I, Martín J, Erdmann J,

Ferrusquía-Acosta J, Garcia-Etxebarria K, Izquierdo-Sanchez L, Bettini LR, Sumoy L, Terranova L, Moreira L, Santoro L, Scudeller L, Mesonero F, Roade L, Rühlemann MC, Schaefer M, Carrabba M, Riveiro-Barciela M, Figuera Basso ME, Valsecchi MG, Hernandez-Tejero M, Acosta-Herrera M, D'Angiò M, Baldini M, Cazzaniga M, Schulzky M, Cecconi M, Wittig M, Ciccarelli M, Rodríguez-Gandía M, Bocciolone M, Miozzo M, Montano N, Braun N, Sacchi N, Martínez N, Özer O, Palmieri O, Faverio P, Preatoni P, Bonfanti P, Omodei P, Tentorio P, Castro P, Rodrigues PM, Blandino Ortiz A, de Cid R, Ferrer R, Gualtierotti R, Nieto R, Goerg S, Badalamenti S, Marsal S, Matullo G, Pelusi S, Juzenas S, Aliberti S, Monzani V, Moreno V, Wesse T, Lenz TL, Pumarola T, Rimoldi V, Bosari S, Albrecht W, Peter W, Romero-Gómez M, D'Amato M, Duga S, Banales JM, Hov JR, Folseraas T, Valenti L, Franke A, Karlsen TH, Severe Covid-19 GWAS Group. Genomewide Association Study of Severe Covid-19 with Respiratory Failure. N Engl J Med [Internet] 2020 Jun 17; PMID:32558485

77. Ton A-T, Gentile F, Hsing M, Ban F, Cherkasov A. Rapid Identification of Potential Inhibitors of SARS-CoV-2 Main Protease by Deep Docking of 1.3 Billion Compounds. Mol Inform [Internet] 2020 Aug;39(8):e2000028. PMID:32162456

78. Abbasi BA, Saraf D, Sharma T, Sinha R, Singh S, Gupta P, Sood S, Gupta A, Rawal K. Identification of vaccine targets & design of vaccine against SARS-CoV-2 coronavirus using computational and deep learning-based approaches. Open Science Framework. 2020.

79. Ong E, Wang H, Wong MU, Seetharaman M, Valdez N, He Y. Vaxign-ML: supervised machine learning reverse vaccinology model for improved prediction of bacterial protective antigens. Bioinformatics [Internet] 2020 May 1;36(10):3185–3191. PMID:32096826

80. Malone B, Simovski B, Moliné C, Cheng J, Gheorghe M, Fontenelle H, Vardaxis I, Tennøe S, Malmberg J-A, Stratford R, Clancy T. Artificial intelligence predicts the immunogenic landscape of SARS-CoV-2: toward universal blueprints for vaccine designs. Bioinformatics. bioRxiv; 2020. p. 536.

81. Keshavarzi Arshadi A, Webb J, Salem M, Cruz E, Calad-Thomson S, Ghadirian N, Collins J, Diez-Cecilia E, Kelly B, Goodarzi H, Yuan JS. Artificial Intelligence for COVID-19 Drug Discovery and Vaccine Development. Front Artif Intell [Internet] 2020 Aug 18;3:10. [doi: 10.3389/frai.2020.00065]

82. Hoffmann M, Kleine-Weber H, Schroeder S, Krüger N, Herrler T, Erichsen S, Schiergens TS, Herrler G, Wu N-H, Nitsche A, Müller MA, Drosten C, Pöhlmann S. SARS-CoV-2 Cell Entry Depends on ACE2 and TMPRSS2 and Is Blocked by a Clinically Proven Protease Inhibitor. Cell [Internet] 2020 Apr 16;181(2):271–280.e8. PMID:32142651

83. Santus E, Li C, Yala A, Peck D, Soomro R, Faridi N, Mamshad I, Tang R, Lanahan CR, Barzilay R, Hughes K. Do Neural Information Extraction Algorithms Generalize Across Institutions? JCO Clin Cancer Inform [Internet] 2019 Jul;3:1–8. PMID:31310566

84. Santus E, Schuster T, Tahmasebi AM, Li C, Yala A, Lanahan CR, Prinsen P, Thompson SF, Coons S, Mynderse L, Barzilay R, Hughes K. Exploiting Rules to Enhance Machine Learning

in Extracting Information From Multi-Institutional Prostate Pathology Reports. JCO Clin Cancer Inform [Internet] 2020 Oct;4:865–874. PMID:33006906

85. Adnan K, Akbar R. An analytical study of information extraction from unstructured and multidimensional big data. J Big Data [Internet] 2019 Dec 17;6(1):1. [doi: 10.1186/s40537-019-0254-8]

86. Domingo-Fernández D, Baksi S, Schultz B, Gadiya Y, Karki R, Raschka T, Ebeling C, Hofmann-Apitius M, Kodamullil AT. COVID-19 Knowledge Graph: a computable, multi-modal, cause-and-effect knowledge model of COVID-19 pathophysiology. Bioinformatics. bioRxiv; 2020. p. 551.

87. Cowen L, Ideker T, Raphael BJ, Sharan R. Network propagation: a universal amplifier of genetic associations. Nat Rev Genet [Internet] 2017 Sep;18(9):551–562. PMID:28607512

88. Humayun F, Domingo-Fernández D, Paul George AA, Hopp M-T, Syllwasschy BF, Detzel MS, Hoyt CT, Hofmann-Apitius M, Imhof D. A Computational Approach for Mapping Heme Biology in the Context of Hemolytic Disorders. Front Bioeng Biotechnol [Internet] 2020 Mar 6;8:74. PMID:32211383

89. Nelson W, Zitnik M, Wang B, Leskovec J, Goldenberg A, Sharan R. To Embed or Not: Network Embedding as a Paradigm in Computational Biology. Front Genet [Internet] 2019 May 1;10:381. PMID:31118945

90. Hamilton WL, Ying R, Leskovec J. Representation Learning on Graphs: Methods and Applications [Internet]. arXiv [csSI]. 2017. Available from: http://arxiv.org/abs/1709.05584

91. Allen Institute For AI. COVID-19 Open Research Dataset Challenge (CORD-19) [Internet]. [cited 2020 Oct 4]. Available from: https://www.kaggle.com/allen-institute-for-ai/CORD-19-research-challenge

92. Ristoski P, Rosati J, Di Noia T, De Leone R, Paulheim H. RDF2Vec: RDF graph embeddings and their applications. Lecue F, editor. SW [Internet] 2019 May 23;10(4):721–752. [doi: 10.3233/SW-180317]

93. COVID-19 SARS-CoV-2 preprints from medRxiv and bioRxiv [Internet]. [cited 2020 Oct 4]. Available from: https://connect.biorxiv.org/relate/content/181

94. The Lancet Public Health. COVID-19 puts societies to the test. The Lancet Public Health [Internet] 2020 May;5(5):e235. [doi: 10.1016/S2468-2667(20)30097-9]

95. Alam F, Dalvi F, Shaar S, Durrani N, Mubarak H, Nikolov A, Da San Martino G, Abdelali A, Sajjad H, Darwish K, Nakov P. Fighting the COVID-19 Infodemic in Social Media: A Holistic Perspective and a Call to Arms [Internet]. arXiv [csIR]. 2020. Available from: http://arxiv.org/abs/2007.07996

96. Song X, Petrak J, Jiang Y, Singh I, Maynard D, Bontcheva K. Classification Aware Neural

Topic Model and its Application on a New COVID-19 Disinformation Corpus [Internet]. arXiv [csLG]. 2020. Available from: http://arxiv.org/abs/2006.03354

97. Li Y, Grandison T, Silveyra P, Douraghy A, Guan X, Kieselbach T, Li C, Zhang H. Jennifer for COVID-19: An NLP-Powered Chatbot Built for the People and by the People to Combat Misinformation. Workshop on Natural Language Processing for COVID-19 (NLP-COVID) 2020.

98. Fogel AL, Kvedar JC. Artificial intelligence powers digital medicine. NPJ Digit Med [Internet] 2018 Mar 14;1:5. PMID:31304291

99. Shen B, Yi X, Sun Y, Bi X, Du J, Zhang C, Quan S, Zhang F, Sun R, Qian L, Ge W, Liu W, Liang S, Chen H, Zhang Y, Li J, Xu J, He Z, Chen B, Wang J, Yan H, Zheng Y, Wang D, Zhu J, Kong Z, Kang Z, Liang X, Ding X, Ruan G, Xiang N, Cai X, Gao H, Li L, Li S, Xiao Q, Lu T, Zhu Y, Liu H, Chen H, Guo T. Proteomic and Metabolomic Characterization of COVID-19 Patient Sera. Cell [Internet] 2020 Jul 9;182(1):59–72.e15. PMID:32492406

100. Ahmed Z, Mohamed K, Zeeshan S, Dong X. Artificial intelligence with multi-functional machine learning platform development for better healthcare and precision medicine. Database [Internet] 2020 Jan 1;2020. PMID:32185396

101. Shah P, Kendall F, Khozin S, Goosen R, Hu J, Laramie J, Ringel M, Schork N. Artificial intelligence and machine learning in clinical development: a translational perspective. NPJ Digit Med [Internet] 2019 Jul 26;2:69. PMID:31372505

102. Mahalle PN, Sable NP, Mahalle NP, Shinde GR. Data Analytics: COVID-19 Prediction Using Multimodal Data. MATHEMATICS & COMPUTER SCIENCE. 2020.

103. Overmyer KA, Shishkova E, Miller IJ, Balnis J, Bernstein MN, Peters-Clarke TM, Meyer JG, Quan Q, Muehlbauer LK, Trujillo EA, He Y, Chopra A, Chieng HC, Tiwari A, Judson MA, Paulson B, Brademan DR, Zhu Y, Serrano LR, Linke V, Drake LA, Adam AP, Schwartz BS, Singer HA, Swanson S, Mosher DF, Stewart R, Coon JJ, Jaitovich A. Large-scale Multi-omic Analysis of COVID-19 Severity. medRxiv [Internet] 2020 Jul 19; PMID:32743614

104. Ray S, Srivastava S. COVID-19 Pandemic: Hopes from Proteomics and Multiomics Research. OMICS [Internet] 2020 Aug;24(8):457–459. PMID:32427517

105. Schork NJ. Artificial Intelligence and Personalized Medicine. In: Von Hoff DD, Han H, editors. Precision Medicine in Cancer Therapy [Internet] Cham: Springer International Publishing; 2019. p. 265–283. [doi: 10.1007/978-3-030-16391-4_11]

106. The Cost of Sequencing a Human Genome [Internet]. [cited 2020 Oct 6]. Available from: https://www.genome.gov/about-genomics/fact-sheets/Sequencing-Human-Genome-cost

107. Dias R, Torkamani A. Artificial intelligence in clinical and genomic diagnostics. Genome Med [Internet] 2019 Nov 19;11(1):70. PMID:31744524

108. Kelly CJ, Karthikesalingam A, Suleyman M, Corrado G, King D. Key challenges for delivering clinical impact with artificial intelligence. BMC Med [Internet] 2019 Oct 29;17(1):195. PMID:31665002

Supplementary Files

Figures

Main strategic areas of research and development for the realization of AI to fight COVID-19: triage, diagnosis and risk prediction; drug repurposing and development; pharmacogenomics and vaccines; mining of medical literature. The four panels report advantages and actionable solutions exhibited by the AI-aided precision medicine approaches surveyed in this work.