



Host-dependent molecular factors mediating SARS-CoV-2 infection to gain clinical insights for developing effective targeted therapy

Gowhar Shafi¹ · Shruti Desai¹ · Krithika Srinivasan¹ · Aarthi Ramesh² · Rupesh Chaturvedi^{3,4} · Mohan Uttarwar¹

Received: 14 October 2020 / Accepted: 9 March 2021 / Published online: 20 March 2021
© The Author(s), under exclusive licence to Springer-Verlag GmbH Germany, part of Springer Nature 2021

Abstract

Coronavirus disease 2019 (COVID-19), a recent viral pandemic that first began in December 2019, in Hunan wildlife market, Wuhan, China. The infection is caused by a coronavirus, SARS-CoV-2 and clinically characterized by common symptoms including fever, dry cough, loss of taste/smell, myalgia and pneumonia in severe cases. With overwhelming spikes in infection and death, its pathogenesis yet remains elusive. Since the infection spread rapidly, its healthcare demands are overwhelming with uncontrollable emergencies. Although laboratory testing and analysis are developing at an enormous pace, the high momentum of severe cases demand more rapid strategies for initial screening and patient stratification. Several molecular biomarkers like C-reactive protein, interleukin-6 (IL6), eosinophils and cytokines, and artificial intelligence (AI) based screening approaches have been developed by various studies to assist this vast medical demand. This review is an attempt to collate the outcomes of such studies, thus highlighting the utility of AI in rapid screening of molecular markers along with chest X-rays and other COVID-19 symptoms to enable faster diagnosis and patient stratification. By doing so, we also found that molecular markers such as C-reactive protein, IL-6 eosinophils, etc. showed significant differences between severe and non-severe cases of COVID-19 patients. CT findings in the lungs also showed different patterns like lung consolidation significantly higher in patients with poor recovery and lung lesions and fibrosis being higher in patients with good recovery. Thus, from these evidences we perceive that an initial rapid screening using integrated AI approach could be a way forward in efficient patient stratification.

Keywords COVID-19 · SARS-CoV-2 · Molecular biomarkers · Multiomics · Artificial intelligence

Introduction

Coronavirus Disease 2019 (COVID-19), a 2019 pandemic, is caused by the SARS-CoV-2, a novel strain of coronavirus. Other respiratory illnesses such as SARS and MERS are also

caused by coronavirus family of viruses (Poutanen 2018). These viruses cause mild to severe respiratory infections, affecting mammals and birds. Between December 2019 and April 2020, COVID-19 infection rose as a rapidly spreading pandemic, affecting 1,119,109 people in 181 different countries in the world (statistics till manuscript submission). While the total death toll increased beyond 58,955 people across different countries, recovery rates stood higher constituting more than 226,873 people (Johns Hopkins University 2020a, b; World Health Organization 2020).

Although highly contagious, only ~20% of the infected needed severe hospitalization. However, the extremely severe cases demand high medical attention and utilize major hospital resources from patient beds, intensive care units (ICU), respiratory support to ventilators. With limited availability of these healthcare resources, the condition turned unmanageable (Grasselli et al. 2020; Rosenbaum 2020).

Communicated by Stefan Hohmann.

✉ Gowhar Shafi
gowhar.shafi@indx.ai

¹ iNDX.Ai, Cupertino, CA, USA

² School of Science, Monash University Malaysia, Jalan Lagoon Selatan, Bandar Sunway, 47500 Subang Jaya, Selangor, Malaysia

³ School of Biotechnology, Jawaharlal Nehru University, New Delhi 110067, India

⁴ Nanofluidiks Pvt. Ltd. Jawaharlal Nehru University-Foundation for Innovation New Delhi, New Delhi 110067, India

While age and symptoms matter in COVID-19 patient care, they are not necessarily accurate predictors of disease severity and deaths. Molecular biomarkers like C-reactive protein, interleukin-6 (IL-6), eosinophils and chest X-ray/computed tomography (CT) findings like lung consolidation, lung lesions and fibrosis have shown significant associations with COVID-19 disease severity as evidenced by several studies (Huang et al. 2020; Yun et al. 2021; Broman et al. 2020; Hashemian et al. 2020).

Since only few of these markers are tested clinically in COVID-19 patients, a robust and high-throughput technique like AI-enabled multiomics can integrate and assist in rapid screening of patients for efficient stratification. This way, valuable information like disease severity can be acquired at a high pace (Kumar et al. 2020).

We at iNDX.Ai, as AI innovators in precision oncology, are also enduring to repurpose our platforms like iCore (image analytics) and iCE (Integrated Correlation Engine—AI-based multiomics investigation software) to COVID-19 investigation so as to achieve rapid utility of enormous patient data, powerful patient stratification and feature prediction. Due to the limited knowledge of the molecular aspects of COVID-19 infection so far, an AI-enabled comprehensive evaluation of host molecular pathways in SARS-CoV-2 replication could also facilitate a better understanding of the disease. This review article thus aims to address the combined benefits of molecular biomarkers

and clinical findings in efficient COVID-19 patient screening and stratification.

Current status of COVID-19

Disease origin/progression mechanisms

The infection was first detected in Wuhan, Hubei province, China, and is believed to have spread from wildlife wet market. The disease is caused by a strain SARS-CoV-2, that is phylogenetically related to severe acute respiratory syndrome-like (SARS-like) bat viruses, raising a possibility of bats to be its primary hosts. The virus after its ingress into human system, incubates and replicates before infecting multiple cells causing serious symptoms like upper respiratory tract infection, fever, sore throat, myalgias, etc. However, some infected cohort are asymptomatic with internally aggravating and advancing disease. A list of all immediate and persistent symptoms are presented in Table 1.

With its rapid spread, COVID-19 has invaded more than a 100 countries across the world with various degrees of illnesses and fatalities. Regarded as pandemic by the World Health Organization (WHO), infections started overwhelming hospitals, healthcare professionals and medical resources (World Health Organization 2020). There are no clinically approved drugs/therapies for COVID-19; however, broad spectrum antibodies are under investigation in clinical trials.

Table 1 Table presents immediate and persistent symptoms of COVID-19 infection

Immediate symptoms				Persistent symptoms (even after recovery)	
Common symptoms	References	Severe symptoms	References	Symptoms	References
Fever or chills	Centers for Disease Control and Prevention (2020)	Breathing trouble/shortness of breath (pneumonia)	Centers for Disease Control and Prevention (2020)	Cough	Ladds et al. (2020)
Cough		Continuous pressure in the chest		Breathlessness	
Fatigue		Inability to wake or stay awake		Fever	
Muscle or body aches		Bluish lips or face		Sore throat	
Headache				Chest pain	
Loss of taste or smell				Palpitations	
Sore throat				Myalgia	
Congestion or runny nose				Neurological symptoms	
Nausea or vomiting				Skin rashes	
Diarrhea				Diarrhea	
Tingling or numbness in hands and feet	Harvard Health Publishing (2021a, b)			Low oxygen saturations	
Delirium and confusions					
Seizures and stroke					

Vaccine investigations are also ongoing with several efforts approved by the US Food and Drug Administration (FDA) (Shereen et al. 2020; The New York Times 2021).

Factors of predisposition and patient severity

While any healthy individual in a population is vulnerable to SARS-CoV-2 infection, several health factors and comorbidities elevate the risk of some individuals for both infections and deaths. Smoking and alcohol can be major lifestyle risk factors of COVID-19 infections. However, there is no large statistical evidence to prove this phenomenon.

In a retrospective multicenter cohort study conducted in affected patients in Wuhan, China, several compulsive comorbidities have been explored. Majority (~30%) of the affected individuals have been identified with hypertension, making it the major susceptible factor. Diabetes was the second most susceptibility factor followed by coronary heart disease which were 19% and 8%, respectively. Old age was also identified to be a significant factor for infection severity and death (Zhou et al. 2020).

Global diagnostic tools

As the virus rapidly progresses to the respiratory tract through 5–7 days of infection, a nasopharyngeal (NP) swab and/or an oropharyngeal (OP) swab is recommended for screening and early diagnosis (Tang et al. 2020). For more severe symptoms, lower respiratory specimens such as sputum and/or endotracheal aspirate or bronchoalveolar lavage is investigated. Blood and stool samples are also considered for additional examinations. Since coronaviruses are enveloped single stranded ribonucleic acid (RNA) viruses, laboratory diagnosis is performed using real-time reverse

transcription polymerase chain reaction (rRT-PCR) reactions (Center for Infectious Disease Research and Policy 2020).

In addition, chest X-rays and CT scans are also regularly captured to diagnose severe respiratory conditions like pneumonia in critically ill patients. Viral genome sequencing is also a developing scenario to monitor any evolving mutations in the virus which, however, has not been widely adapted for clinical investigation (World Health Organization 2020; ScienceDaily 2020).

Treatment/control methods

As per Centers for Disease Control and Prevention (CDC), there are no current preventive or treatment drugs approved by FDA. However, several drugs and strategies are being utilized for palliative care. As a measure for supportive care, oxygen supplementation and mechanical ventilation are being utilized in intensive care facilities for COVID-19 infection (Centers for Disease Control and Prevention 2020).

Several broad-spectrum antiviral drugs such as remdesivir are being extensively investigated for treating COVID-19 infected patients. Certain antimalarial and anti-inflammatory drugs such as hydroxychloroquine and chloroquine are also under clinical trials for use as prophylactic and therapeutic for COVID-19 infections (Centers for Disease Control and Prevention 2020). A list of commonly utilized medications for management of COVID-19 patients is presented in Table 2. Precautions, strict quarantining and patient isolation are also effective strategies followed for containing the viral spread and transmission (Centers for Disease Control and Prevention 2020).

Table 2 Table presents a list of COVID-19 primary management medications, drug class and treatment indications

Medication	Drug class	Indications	References
Corticosteroids (Dexamethasone, Prednisone, Methylprednisolone)	Anti-inflammatory	Recommended during hyper-immune reactions/cytokine storm upon viral infection that damages lungs and other organs	Harvard Health Publishing (2021a, b)
Remdesivir	Anti-viral	Recommended against viral replication	Harvard Health Publishing (2021a, b)
Anticoagulation drugs (“blood thinners”)	Anti-coagulant	Recommended to prevent/treat blood clots	Harvard Health Publishing (2021a, b)
Bamlanivimab, Casirivimab and Imdevimab)	Monoclonal antibodies	Recommended against viral attachment to host cells (action against coronavirus’s spike protein)	Harvard Health Publishing (2021a, b)
Convalescent plasma	Antibody therapy	Antibodies from recovered patients recommended to act against replication and spread of virus	Harvard Health Publishing (2021a, b)

Challenges in clinical settings during viral pandemics

Viral pandemics have a history of invading and destroying millions of people across the world since the very known “Spanish flu pandemic” in 1918. Death rates recorded back then were extremely high across the globe with various contributing factors like availability and access to healthcare facilities, public health infrastructure, population density, nutritional factors, etc.

Apart from frontline medical resources, intense research on viral characteristics, drug/vaccine discovery in limited time span and large scale production of the discovered vaccines are all equally challenging. However, economic and technological developments are powerful resources against evolving global emergencies (Oshitani et al. 2008).

The recent COVID-19 pandemic has been a clinically challenging threat to more than 100 countries, with Europe, North America and China hit the hardest. Novel viruses such as COVID-19 display continuously evolving genomic patterns and sequences, which could also vary among different races or ethnicities. In such cases, treatment/management decisions cannot depend only upon clinical manifestations. This contributes to the challenge as their sequences and/or protein structures may not be available for deeper investigations (Li et al. 2020). Several retrospective analyses of integrated patient data including deoxyribonucleic acid (DNA), RNA, host immune profile, cytokine profile, blood and biochemical reports yield direct correlations with patient recovery, disease severity and even death (Thevarajan et al. 2020). Several other technological developments such as virtual reality, AI, AI-assisted robots, AI-assisted chatbots, remote and virtual diagnosis and consultation, etc., also aim to assist healthcare professionals with more rapid and accurate solutions (BROOKINGS 2020; The British Broadcasting Corporation 2020).

This article is one such endeavor to emphasize existing clinical challenges and the routes through which the high throughput AI-enabled multiomics is beginning to assist healthcare from disease prediction, patient stratification to drug/vaccine discovery in COVID-19 pandemic crisis.

Role of AI-enabled multiomics in COVID-19 investigation

The aggressive nature of COVID-19 has led to its intractable spread across various countries rapidly increasing the death rates. Due to its unpredictable complexity and novelty, clinical diagnosis and treatment are burdensome for healthcare professionals.

With the rapidly progressing nature of the virus, a high-throughput approach to detect physical and molecular characteristics more rapidly is essential to keep up diagnostic accuracy and enable drug/vaccine development. Thus several AI-enabled platforms are spiking into COVID-19 investigation from diagnosis to discovery (UC San Diego Health 2020). With better algorithms and quicker analysis, faster patient care can be clinically adopted. In this section, we discuss some of the AI-enabled tools and their usefulness in COVID-19 interventions.

AI-enabled multiomics and patient stratification

The SARS-CoV-2 has been identified to cause infections irrespective of an individual’s current or previous medical conditions. However, multiple host-dependent molecular factors contribute to recovery/serious illness in addition to clinical manifestations. Clinically, most of the infected patients have been documented to present varying degrees of fever, cough, fatigue, myalgias, shortness of breath, to name a few. However, some are physically asymptomatic with or without internal advancement as mentioned earlier (Qiu et al. 2020; Pan et al. 2020; World Health Organization 2020). This suggests that clinical presentations alone cannot be predictive of virus penetrance and progression internally.

For instance, a case study published in 2020 by Thevarajan et al., utilized multiomics approach to investigate host immune responses of a 47-year-old female COVID-19 patient. Upon retrospective investigations at different time points throughout her infection, several factors like antibody secreting cells (ASCs), Follicular helper T cells (T_{FH} cells), activated CD4 + T cells, CD8 + T cells, immunoglobulin M (IgM) and immunoglobulin G (IgG) antibodies were found to be elevated. Upon cytokine and chemokine profiling during symptomatic phase of the disease, decreased pro-inflammatory cytokines and chemokines were observed. A risk genotype in Interferon Induced Transmembrane Protein 3 gene (*IFITM3* gene) (rs12252) for influenza was also observed on genetic testing. Interestingly, 26% of the Chinese population carried the risk genotype as per 1000 Genomes Project which was surmised to reflect COVID-19 spectrum in china (Thevarajan et al. 2020). This case study illustrates the diverse spectrum of essential host-dependent molecular parameters that could contribute to patient recovery in COVID-19 disease. However, healthcare professionals are needed to be focused on treating healthcare needs of patients. Thus AI-enabled multiomic models could efficiently stratify patients based on strictly trained algorithms promoting appropriate care, faster recovery and proper allocation of hospital resources.

Polymorphisms in genes like Angiotensin I Converting Enzyme 2 (ACE2) are also contributing to increased vulnerability of COVID-19 infection (Alimadadi et al. 2020).

AI-based algorithms can be trained with large genomic datasets and can be utilized as powerful tools to screen huge population for polymorphisms at high-throughput. Neural network based algorithms can also be utilized for large-scale screening and prediction of COVID-19 infections using complex data like patient respiratory patterns (Alimadadi et al. 2020).

Other significant molecular stratifiers of COVID-19 patients include immune markers, metabolome (e.g., liver enzyme alanine aminotransferase), cell surface receptors, genes responsible for other comorbidities such as cardiac diseases, diabetes and so on (Jiang et al. 2020). IL-6 is one such molecular marker, which has been identified to significantly increase with the disease severity among COVID-19 infected cohort and are also strongly associated with the need for mechanical ventilation. Increasing levels of this marker also predicted respiratory failure with high accuracy as shown in studies (Herold et al. 2020). Figure 1 depicts the levels of IL-6 in severely affected patients as compared with non-severe cases among COVID-19 infected patients investigated from 9 studies.

In addition to the clinical and molecular biomarkers, AI has been intelligently utilized by researchers in Massachusetts Institute of Technology (MIT) to predict cough patterns of asymptomatic COVID-19 patients. The study utilizes thousands of cough patterns to develop an AI model that

can discriminate cough of an asymptomatic individual from uninfected individuals. If approved by FDA, this AI-based model can be an efficient tool for asymptomatic patient stratification from healthy individuals thereby assisting in reduction of viral spread and infections (MIT News 2020).

Patient surveillance and tracking, virtual healthcare assistance and high-speed data capture and retrieval

Migration and community gathering are the major causes of rapid multiplication of pandemic infections. Since, SARS-CoV-2 is capable of incubating asymptomatic for ~7–14 days inside the human system, early isolation before clinical presentation has been challenging. In such cases, AI acts beneficial to monitor migration in populations at risk or affected so as to enable efficient tracking. AI-enabled algorithms can also accurately predict individual risk of infection with sufficient information on family history, lifestyle, social and travel history, etc., AI models are also trained as information bots to deliver reliable general information about infection, speed of transmission, recovery rates, etc., catering general awareness (Panch et al. 2019; Kuziemyk et al. 2019).

Multiple AI-based data dashboards are also functional and beneficial in tracking infected cases, recoveries and

High interleukin-6 levels consistent in severe COVID-19 infections as shown in literature

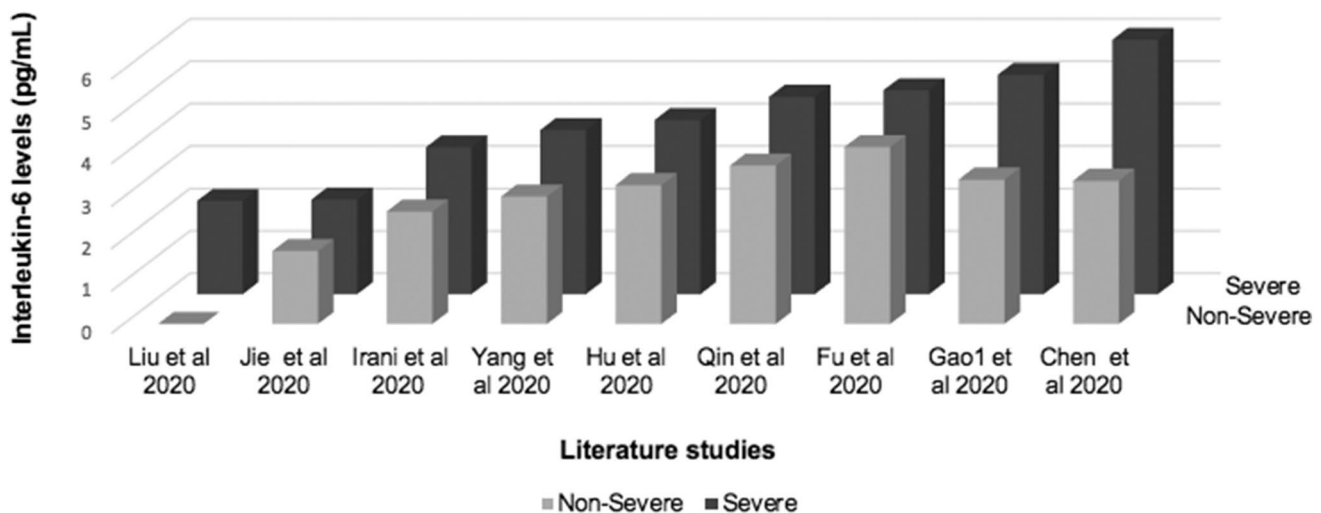


Fig. 1 Relationship between IL-6 and COVID-19 disease severity. IL-6 levels are reported to be higher in severe COVID-19 infections than non-severe patients in studies mentioned below. Figure 1 is a compiled reanalysis of IL-6 data from 7 studies which consistently

showed increased levels of IL-6 in severe cases than non-severe ones (Chen et al. 2020; Gao et al. 2020; Liu et al. 2020a, b; Yang et al. 2020; Qin et al. 2020; Fu et al. 2020; Thevarajan et al. 2020; Wu et al. 2020; Wang et al. 2020)

deaths across the globe. They also display interactive maps for better visualization. Some of such platforms include Johns Hopkins' CSSE and Microsoft Bing's COVID-19 tracker (Naudé 2020).

Computer based algorithms are greatly helpful in these situations accommodating not only immediate health-care needs but also catering to general information, live tracking, disease screening, etc., for future disease-based interventions.

AI-enabled multiomics in disease severity, clinical trial recruitment and drug development studies

Age has been considered a major risk factor of COVID-19 infection based on worldwide statistics. A model-based investigation of age group vs case fatality ratio also rendered a linear relationship of 0.32% vs 6.4% in individuals older than 60 years of age (Verity et al. 2020). However, infection and disease severity ratios may not necessarily follow age relationships, as the infections have been recorded in younger population as well (Verity et al. 2020).

Recently, an AI-based prediction of COVID-19 severity, investigated by NYU Grossman School of Medicine and the Courant Institute of Mathematical Sciences at New York University, in partnership with Wenzhou Central Hospital and Cangnan People's Hospital, both in Wenzhou, China, was able to predict the need for patient respiratory support based on molecular factors. Three major features were utilized in the study which includes liver enzyme alanine aminotransferase (ALT) levels, myalgia and hemoglobin levels reaching an accuracy of up to 80% (ScienceDaily 2020).

Similarly, another study by Tan et al., identified significantly lower lymphocyte percentages in patients with severe and fatal disease using time-lymphocyte percentage model (TLM) (Tan et al. 2020). The study also identified that the patients who recovered from infections showed an increasing trend of lymphocyte percentage making them a useful predictor of disease severity.

Mortality also hits some patient cohorts who present extremely elevated cytokine profiles. From a retrospective study in China, it has been identified that some infected patients presented excessively high levels of ferritin and IL-6 regarded as "cytokine storm syndrome", being interpreted as a cause of mortality. Virus induced hyperinflammation resulting in increased levels of interleukin-2 (IL-2), interleukin-7 (IL-7), granulocyte colony stimulating factor, interferon- γ -inducible protein-10, macrophage inflammatory protein (MIP)-1alpha and tumour necrosis factor alpha could all be responsible for elevating patient fatality. Application of AI-enabled multiomics analysis in these areas could accurately predict the need for immunosuppressive drugs which could otherwise be life-threatening in viral infections (Mehta et al. 2020). A compiled representation of multiple

molecular markers from over 450 individuals accumulated and reanalyzed from different studies and case reports on COVID-19 is depicted in Fig. 2 (Huang et al. 2020; Broman et al. 2020; Hashemian et al. 2020). This shows the contribution of molecular representatives in determining potential candidates for clinical trial and/or drug development investigations.

In addition to patient stratification and clinical trial recruitment, AI-based drug prediction models are highly evolving. With the advancing availability of enormous patient data and viral structural/sequential information, it is achievable to train AI algorithms to predict efficacy against certain class/specific drugs. For instance, an AI-based model developed by researchers in Korea and United States has been identified to predict targeted drugs to inhibit the replication of 2019-nCoV. The model utilizes 1 dimensional (1D) string of input sequences as against the unavailability of 3 dimensional (3D) protein structures, predicting the inhibitory potencies of antiretroviral drugs like atazanavir, efavirenz, retonavir and dolutegravir using K_p , K_d and IC50 values. Molecular docking studies have also been utilized for deeper investigation of prediction accuracies (Beck et al. 2020). Due to the flexibility of AI-enabled models, they show great potential in several drug developmental studies.

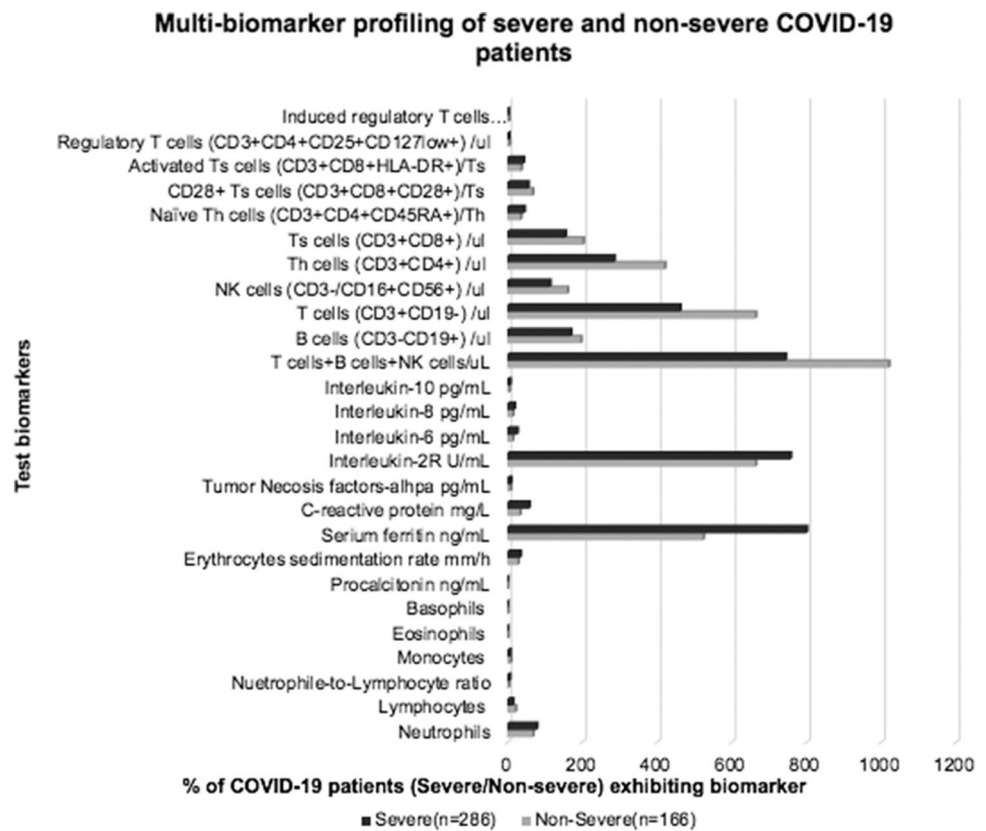
AI-enabled high throughput prediction of clinical images for early signs of pneumonia

Pneumonia is a major cause of severity and death in COVID-19 infections. However, even expert radiologists are overwhelmed with huge radiomic data with increasing numbers of CT scans and chest X-rays of affected individuals. Thus, regular monitoring of images for early diagnosis is a challenge due to continuously increasing burden of infections.

UC San Diego Health has developed an AI-assisted model enabled by Amazon Web Services (AWS) to accurately predict, monitor and diagnose pneumonia with patient CT scans and chest X-ray images. This model has provided experts with new insights in about 2000 clinical images. Interestingly it has also predicted a patient's risk of developing pneumonia with early signs in chest X-ray, which would have been missed otherwise (ScienceDaily 2020).

Several other AI algorithms are being investigated for image analytics in COVID-19 patients. These algorithms are usually trained with equal/sufficient data under 3 major categories including chest X-rays from healthy individuals, pneumonia and COVID-19 affected patients. Training is usually focused to maintain consistent prediction accuracy even with low quality data with appropriate preprocessing. Radiomic features are then extracted, balanced and trained under various data classifiers such as Random Forest (RF), Decision Trees (DT), etc. Such models have established a maximum accuracy of 0.973 using RF (Kumar et al. 2020).

Fig. 2 Multi-biomarker profiling of severe and non-severe COVID-19 patients. Figure represents the levels of multiple biomarkers as observed in severe and non-severe COVID-19 patients. Cells mediating immune reactions including stimulatory T cells, helper T cells, NK cells, T cells and B cells are higher in non-severe cases indicating events of immune reactions. On the other hand, biomarkers such as C-reactive proteins, serum ferritin, neutrophils are all higher in severe COVID-19 patients. These individual biomarkers alone may not be conclusive of the disease, but their initial screening may enable better understanding of the disease severity for efficient patient stratification



AI-enabled multiomics studies, apart from the early prediction of pneumonia, can also be utilized to extract various other useful information for treatment surveillance and monitoring, which includes pulmonary consolidation, disease resolution and fibrosis from chest X-rays and CT scans. This information at a higher accuracy and faster pace can assist timely stratification and resource allocation (Castiglioni et al. 2020).

A compiled representation of chest CT findings from severe and non-severe COVID-19 patients is depicted in Fig. 3. The figure shows significantly high prevalence of ground glass opacities and lung consolidation, fibrosis and air bronchogram in severe COVID-19 patients as compared with non-severe counterparts. Figure 2 also represents an increased prevalence of these measurement parameters during disease progression which dropped during the phase of recovery (Yun et al. 2021).

Though developers claim further investigation for AI-assisted prediction models, efforts like this are much required for high-speed clinical interventions before witnessing serious health effects.

AI-enabled COVID-19 vaccine investigation

The enormous predictive strength of AI has also been reflected as a potential aid in vaccine development for

COVID-19. Many tools are being developed to thus enhance the laboratory efforts of vaccine development. The “AlphaFold”, a 3D structure prediction tool for genetic sequences, developed by Google DeepMind has predicted the structure and spike of SARS-CoV-2 enabling scientists to better understand their receptor binding and cellular entry. A researcher from the University of Texas at Austin, has also created a 3D atomic scale map to predict the spike protein. Many other efforts are also underway in this field so as to assist in the development of anti-viral components and vaccines (Bullock et al. 2020).

Another notable contribution by Inovio pharmaceuticals, in the generation of DNA-based vaccine called INO-4800. The technique has utilized viral genetic sequences to train the machine learning algorithm (Thanh et al. 2020). Similar approaches by Moderna therapeutics generated a modified messenger RNA (mRNA)-based vaccine called mRNA-1273 (Thanh et al. 2020), which has been currently approved in Switzerland and for emergency use in U.S., U.K., E.U., others (The New York Times 2021). Several other vaccine efforts like Comirnaty (Pfizer-BioNTech), AZD1222 (an Oxford-AstraZeneca effort) which are currently in clinical phase trial 2 and 3 have been approved by FDA for use in several countries (Information dated February, 2021) listed in Table 3 (The New York Times 2021). These approved vaccines were developed based on many strategies such as

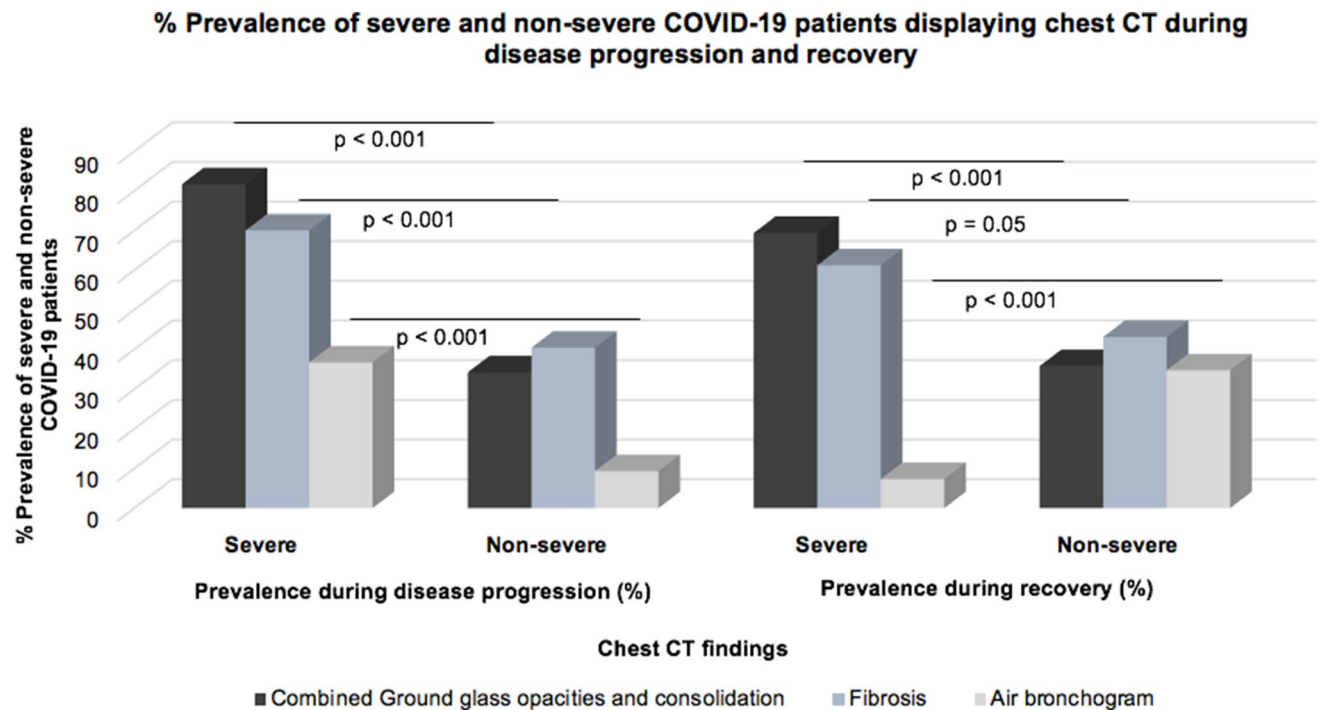


Fig. 3 A retrospective investigation of chest CT findings to assist in COVID-19 patient stratification of severe and non-severe disease pattern. The chest CT findings presented in the figure shows 2 vital findings. The first major finding is that the prevalence of COVID-19 patients exhibiting conditions like ground glass opacities, lung consolidation, fibrosis and air bronchogram were higher during disease progression which dropped upon recovery phase of the disease. The

2nd major finding is that all the parameters (ground glass opacities, lung consolidation, fibrosis and air bronchogram) were higher in patients with severe disease pattern than non-severe disease pattern. From these evidences, an initial screening of COVID-19 patients using chest CT could be a beneficial strategy towards patient stratification

Table 3 Table presents a list of COVID-19 vaccines, method of development and clinical approval status

Vaccine	Method of development	Status of clinical trial	Status of approval	Efficacy	Country of origin	References
Pfizer-BioNTech	mRNA	2, 3	Approved and emergency use	95%	United States and Germany	The New York Times (2021)
Moderna	mRNA	3	Approved and emergency use	94.5%	United States	
Gamaleya	Ad26, Ad5	3	Early and emergency use	91.6%	Russia	
Oxford-AstraZeneca	ChAdOx1	2, 3	Emergency use	82.4%	United Kingdom and Sweden	
CanSino	Ad5	3	Limited use	91.6%	China	Logunov et al. (2021)
Johnson & Johnson	Ad26	3	Not approved yet	72% in United States	United States	The New York Times (2021)
Vector Institute	Protein	3	Early use	100%	Russia	Thomson Reuters (2021)
Novavax	Protein	3	Not yet approved	89.3%	United States	The New York Times (2021)
Sinopharm	Inactivated	3	Approved and emergency use	79.34%	China	
Sinovac	Inactivated	3	Conditional approval and emergency use	50.38%	China	
Bharat Biotech	Inactivated	3	Emergency use	Unknown	India	

mRNA, protein, inactivated viruses, etc. (The New York Times 2021). Many of these vaccines have utilized AI-enabled systems for initial screening for vaccine development to pace up the whole process of laboratory investigations and clinical trials thus developing solutions faster than expected.

Viral epidemics or pandemics usually experience multiple evolutionary stages with its nuclear material thus developing newer strains. Sometimes these new strains may be more infectious even than the original counterparts (Johns Hopkins University 2020a, b). These strains are a common pattern in any viral or bacterial outbreak before the invading organisms reach the state of non-infectiousness (Johns Hopkins University 2020a, b). Similarly, the 2019 pandemic also gave rise to newer strains of SARS-CoV-2 coronavirus detected in many parts of the world (Johns Hopkins University 2020a, b). A particular strain called B.1.1.7 was detected in southeastern England in September 2020 which soon became common in the United Kingdom, accounting for more than 60% of the new COVID-19 cases in December. Several other variants were also detected in areas including South Africa, Brazil, California and others. Unfortunately, several of these strains also alter the composition of the spike protein making them “stickier” (Johns Hopkins University 2020a, b).

More evidences are still required to understand the severity of the disease due to these new strains and most importantly, the efficacy of the current vaccines (Johns Hopkins University 2020a, b). However, many vaccine developers claim that some of these developed vaccines (e.g., mRNA vaccines) could still offer protection against the changing viral pattern and the spike protein but more evidences are still warranted (Johns Hopkins University 2020a, b).

Conclusions and future implications: AI-enabled multilayer data analysis to assist COVID-19 and other global pandemics in the future

Some significant molecular markers and AI-enabled patient tracking, screening and disease detection approaches were discussed in the study. Each of these markers individually and together contribute to efficient patient stratification, prediction of disease severity and therapeutic interventions of COVID-19 as discussed in various sections of the article. With the use of integrated AI-enabled multilayer models, it could be possible to generate high-throughput outcomes from various data sources like clinical, pathological, DNA, RNA, proteome, metabolome, etc. With large data being accumulated over time in different COVID-19 patients and conditions, effective set of training instructions can be delivered to render more accurate predictions of unknown future events (Grapov



Fig. 4 Applications of multiomics in COVID-19 investigation. The figure depicts various applications of multiomics strategies (e.g., genomics, transcriptomics) such as disease susceptibility/risk prediction, disease diagnosis and prognosis, drug response and monitoring. Thus, such strategies could collectively render assistance in the overwhelming healthcare needs during pandemics and beyond

et al. 2018). Since AI-assisted models function with minimal manual interventions, handling large cohorts under stressful circumstances becomes highly feasible (Davenport and Kalakota 2019; Rong et al. 2020; Azuaje 2019). Additionally, AI-enabled multilayer models can relieve biologists of complex data interpretations with the use of simpler visualizations for effective medical decisions (Chakraborty et al. 2017; Chin-Yee and Upshur 2019).

We at iNDX.Ai, are keenly enduring to investigate AI-enabled screening and prediction methods of COVID-19 infections for patient stratification, image analytics and drug discovery using our AI platforms namely iCE and iCore.

We signify that several such multiomics efforts are collectively required to overcome the challenges of COVID-19 in treatment, long-term prevention, prognosis and response prediction, risk analysis, surveillance, etc. which are also presented in Fig. 4. Thus, this article plays its substantial role in emphasizing the competencies of molecular biomarkers and AI-enabled multilayer model in COVID-19 pandemic interventions.

Author contributions GS: contributed to generation of figures, manuscript design and revision; SD: contributed to manuscript design and revision; KS: contributed to manuscript design and figure generation; AR: contributed to manuscript writing and revision; RC: contributed to manuscript revision and data compilation; MU: contributed to manuscript design and revision.

Availability of data and material Will be disclosed upon acceptance.

Code availability Will be disclosed upon acceptance.

Declarations

Conflict of interest The authors declare that they have no conflict of interest.

References

- Alimadadi A, Aryal S, Manandhar I, Munroe PB, Joe B, Cheng X (2020) Artificial intelligence and machine learning to fight COVID-19. *Physiol Genom* 52:200–202. <https://doi.org/10.1152/physiolgenomics.00029.2020>
- Azuaje F (2019) Artificial intelligence for precision oncology: beyond patient stratification. *NPJ Precis Oncol* 3:1–5. <https://doi.org/10.1038/s41698-019-0078-1>
- Beck BR, Shin B, Choi Y, Park S, Kang K (2020) Predicting commercially available antiviral drugs that may act on the novel coronavirus (2019-nCoV), Wuhan, China through a drug-target interaction deep learning model. *bioRxiv*. <https://doi.org/10.1101/2020.01.31.929547>
- Broman N, Rantasärkkä K, Feuth T, Valtonen M, Waris M, Hohenthal U, Rintala E, Karlsson A, Marttila H, Peltola V, Vuorinen T (2020) IL-6 and other biomarkers as predictors of severity in COVID-19. *Ann Med*. <https://doi.org/10.1080/07853890.2020.1840621>
- BROOKINGS (2020) <https://www.brookings.edu/>. Accessed 2020
- Bullock J, Pham KH, Lam CSN, Luengo-Oroz M (2020) Mapping the landscape of artificial intelligence applications against COVID-19. *arXiv*. [arXiv.2003.11336](https://arxiv.org/abs/2003.11336)
- Castiglioni I, Ippolito D, Interlenghi M, Monti CB, Salvatore C, Schiaffino S, Polidori A, Gandola D, Messa C, Sardanelli F (2020) Artificial intelligence applied on chest X-ray can aid in the diagnosis of COVID-19 infection: a first experience from Lombardy, Italy. *medRxiv*. <https://doi.org/10.1101/2020.04.08.20040907>
- Center for Infectious Disease Research and Policy (2020) About COVID-19. <https://www.cidrap.umn.edu/covid-19/>. Accessed 2020
- Centers for Disease Control and Prevention (2020) Coronavirus (COVID-19). <https://www.cdc.gov/coronavirus/>
- Chakraborty I, Choudhury A, Banerjee TS (2017) Artificial intelligence in biological data. *J Inf Technol Softw Eng* 7:2. <https://doi.org/10.4172/2165-7866.1000207>
- Chen X, Zhao B, Qu Y, Chen Y, Xiong J, Feng Y, Men D, Huang Q, Liu Y, Yang B, Ding J (2020) Detectable serum SARS-CoV-2 viral load (RNAemia) is closely associated with drastically elevated interleukin 6 (IL-6) level in critically ill COVID-19 patients. *Clin Infect Dis*. <https://doi.org/10.1038/s41698-019-0078-1>
- Chin-Yee B, Upshur R (2019) Three problems with big data and artificial intelligence in medicine. *Perspect Biol Med* 62:237–256. <https://doi.org/10.1353/pbm.2019.0012>
- Davenport T, Kalakota R (2019) The potential for artificial intelligence in healthcare. *Future Healthc J* 6:94–98. <https://doi.org/10.7861/futurehosp.6-2-94>
- Fu S, Fu X, Song Y, Li M, Pan PH, Tang T, Zhang C, Jiang T, Tan D, Fan X, Sha X (2020) Virologic and clinical characteristics for prognosis of severe COVID-19: a retrospective observational study in Wuhan, China. *medRxiv*. <https://doi.org/10.1101/2020.04.03.20051763>
- Gao Y, Li T, Han M, Li X, Wu D, Xu Y, Zhu Y, Liu Y, Wang X, Wang L (2020) Diagnostic utility of clinical laboratory data determinations for patients with the severe COVID-19. *J Med Virol*. <https://doi.org/10.1002/jmv.25770>
- Grapov D, Fahrman J, Wanichthanarak K, Khoomrung S (2018) Rise of deep learning for genomic, proteomic, and metabolomic data integration in precision medicine. *Omics* 22:630–636. <https://doi.org/10.1089/omi.2018.0097>
- Grasselli G, Pesenti A, Cecconi M (2020) Critical care utilization for the COVID-19 outbreak in Lombardy, Italy: early experience and forecast during an emergency response. *JAMA* 323:1545–1546. <https://doi.org/10.1001/jama.2020.4031>
- Harvard Health Publishing (2021a) <https://www.health.harvard.edu/diseases-and-conditions/covid-19-basics>. Accessed 2021
- Harvard Health Publishing (2021b). <https://www.health.harvard.edu/diseases-and-conditions/treatments-for-covid-19>. Accessed 2021
- Hashemian SM, Shafiq N, Afzal G, Jamaati H, Tabarsi P, Marjani M, Malekmohammad M, Mortazavi SM, Khoundabi B, Mansouri D, Moniri A (2020) Plasmapheresis reduces cytokine and immune cell levels in COVID-19 patients with acute respiratory distress syndrome (ARDS). *Pulmonology*. <https://doi.org/10.1016/j.pulmoe.2020.10.017>
- Herold T, Jurinovic V, Arnreich C, Hellmuth JC, von Bergwelt-Baildon M, Klein M, Weinberger T (2020) Level of IL-6 predicts respiratory failure in hospitalized symptomatic COVID-19 patients. *medRxiv*. <https://doi.org/10.1101/2020.04.01.20047381>
- Huang H, Cai S, Li Y, Li Y, Fan Y, Li L, Lei C, Tang X, Hu F, Li F, Deng X (2020) Prognostic factors for COVID-19 pneumonia progression to severe symptoms based on the earlier clinical features: a retrospective analysis. *medRxiv*. <https://doi.org/10.1101/2020.03.28.20045989>
- Massachusetts Institute of Technology (MIT News) (2020) <https://news.mit.edu/2020/covid-19-cough-cellphone-detection-1029>. Accessed 2020
- Jiang X, Coffee M, Bari A, Wang J, Jiang X, Huang J, Shi J, Dai J, Cai J, Zhang T, Wu Z (2020) Towards an artificial intelligence framework for data-driven prediction of coronavirus clinical severity. *Comput Mater Contin* 63:537–551. <https://doi.org/10.32604/cmc.2020.010691>
- Johns Hopkins University (2020a) COVID-19 INFORMATION. <https://hub.jhu.edu/novel-coronavirus-information/>. Accessed 2020
- Johns Hopkins University (2020b). <https://www.hopkinsmedicine.org/health/conditions-and-diseases/coronavirus/a-new-strain-of-coronavirus-what-you-should-know>. Accessed 2020
- Kumar R, Arora R, Bansal V, Sahayashela VJ, Buechchash H, Imran J, Narayanan N, Pandian GN, Raman B (2020) Accurate prediction of COVID-19 using chest X-ray images through deep feature learning model with SMOTE and machine learning classifiers. *medRxiv*. <https://doi.org/10.1101/2020.04.13.20063461>
- Kuziemy C, Maeder AJ, John O, Gogia SB, Basu A, Meher S, Ito M (2019) Role of artificial intelligence within the telehealth domain. *Yearb Med Inform* 28:035–040. <https://doi.org/10.1055/s-0039-1677897>
- Ladds E, Rushforth A, Wieringa S, Taylor S, Rayner C, Husain L, Greenhalgh T (2020) Persistent symptoms after Covid-19: qualitative study of 114 “long Covid” patients and draft quality principles for services. *BMC Health Serv Res* 20:1–13. <https://doi.org/10.1186/s12913-020-06001-y>
- Li Y, Liu B, Cui J, Wang Z, Shen Y, Xu Y, Yao K (2020) Similarities and evolutionary relationships of COVID-19 and related viruses. *arXiv*. <https://doi.org/10.20944/preprints202003.0316.v1>
- Liu J, Ouyang L, Guo P, Wu SH, Fu P, Chen LY, Yang D, Han YX, Cao KY, Alwalid O, Tao J (2020a) Epidemiological, clinical characteristics and outcome of medical staff infected with COVID-19 in Wuhan, China: a retrospective case series analysis. *medRxiv*. <https://doi.org/10.1101/2020.03.09.20033118>

- Liu L, Gao JY, Hu WM (2020b) Clinical characteristics of 51 patients discharged from hospital with COVID-19 in Chongqing, China. medRxiv. <https://doi.org/10.1101/2020.02.20.20025536>
- Logunov DY, Dolzhenkova IV, Shcheblyakov DV, Tikhvatulin AI, Zubkova OV, Dzharullaeva AS, Kovyshina AV, Lubenets NL, Grousova DM, Erokhova AS, Botikov AG (2021) Safety and efficacy of an rAd26 and rAd5 vector-based heterologous prime-boost COVID-19 vaccine: an interim analysis of a randomised controlled phase 3 trial in Russia. *Lancet*. [https://doi.org/10.1016/S0140-6736\(21\)00234-8](https://doi.org/10.1016/S0140-6736(21)00234-8)
- Mehta P, McAuley DF, Brown M, Sanchez E, Tattersall RS, Manson JJ (2020) COVID-19: consider cytokine storm syndromes and immunosuppression. *Lancet* 395:1033–1034. [https://doi.org/10.1016/S0140-6736\(20\)30628-0](https://doi.org/10.1016/S0140-6736(20)30628-0)
- Naudé W (2020) Artificial Intelligence against COVID-19: an early review. <https://hdl.handle.net/10419/216422>
- Oshitani H, Kamigaki T, Suzuki A (2008) Major issues and challenges of influenza pandemic preparedness in developing countries. *Emerg Infect Dis* 14:875–880. <https://doi.org/10.3201/eid1406.070839>
- Pan L, Mu M, Ren HG, Yang P, Sun Y, Wang R (2020) Clinical characteristics of COVID-19 patients with digestive symptoms in Hubei, China: a descriptive, cross-sectional, multicenter study. *Am J Gastroenterol* 115:766–773. <https://doi.org/10.14309/ajg.0000000000000620>
- Panch T, Mattie H, Celi LA (2019) The “inconvenient truth” about AI in healthcare. *NPJ Digit Med* 2:1–3. <https://doi.org/10.1038/s41746-019-0155-4>
- Poutanen SM (2018) Human coronaviruses. Principles and practice of pediatric infectious diseases, pp 1117–1120. <https://doi.org/10.1016/B978-1-4377-2702-9.00224-5>
- Qin C, Zhou L, Hu Z, Zhang S, Yang S, Tao Y, Xie C, Ma K, Shang K, Wang W, Tian DS (2020) Dysregulation of immune response in patients with COVID-19 in Wuhan, China. *Clin Infect Dis*. <https://doi.org/10.1093/cid/ciaa248>
- Qiu H, Wu J, Hong L, Luo Y, Song Q, Chen D (2020) Clinical and epidemiological features of 36 children with coronavirus disease 2019 (COVID-19) in Zhejiang, China: an observational cohort study. *Lancet Infect Dis* 3099:30198–30205. [https://doi.org/10.1016/S1473-3099\(20\)30198-5](https://doi.org/10.1016/S1473-3099(20)30198-5)
- Thomson Reuters (2021) <https://www.reuters.com/article/us-health-coronavirus-russia-vaccine-vec-idUSKBN290151>. Accessed 2021
- Rong G, Mendez A, Assi EB, Zhao B, Sawan M (2020) Artificial intelligence in healthcare: review and prediction case studies. *Engineering* 6:291–301. <https://doi.org/10.1016/j.eng.2019.08.015>
- Rosenbaum L (2020) Facing COVID-19 in Italy—ethics, logistics, and therapeutics on the epidemic’s front line. *N Engl J Med*. <https://doi.org/10.1056/NEJMp2005492>
- ScienceDaily (2020) <https://www.sciencedaily.com/>. Accessed 2020
- Shereen MA, Khan S, Kazmi A, Bashir N, Siddique R (2020) COVID-19 infection: origin, transmission, and characteristics of human coronaviruses. *J Adv Res* 24:91–98. <https://doi.org/10.1016/j.jare.2020.03.005>
- Tan L, Wang Q, Zhang D, Ding J, Huang Q, Tang YQ, Wang Q, Miao H (2020) Lymphopenia predicts disease severity of COVID-19: a descriptive and predictive study. *Signal Transduct Target Ther* 5:1–3. <https://doi.org/10.1038/s41392-020-0148-4>
- Tang YW, Schmitz JE, Persing DH, Stratton CW (2020) The laboratory diagnosis of COVID-19 infection: current issues and challenges. *J Clin Microbiol*. <https://doi.org/10.1128/JCM.00512-20>
- Thanh LT, Andreadakis Z, Kumar A, Gómez RR, Tollefsen S, Saville M, Mayhew S (2020) The COVID-19 vaccine development landscape. *Nat Rev Drug Discov* 19:305–306. <https://doi.org/10.1038/d41573-020-00073-5>
- The British Broadcasting Corporation (2020) <https://www.bbc.com/news/>. Accessed 2020
- The New York Times (2021) <https://www.nytimes.com/interactive/2020/science/coronavirus-vaccine-tracker.html#novavax>. Accessed 2021
- Thevarajan I, Nguyen TH, Koutsakos M, Druce J, Caly L, van de Sandt CE, Jia X, Nicholson S, Catton M, Cowie B, Tong SY (2020) Breadth of concomitant immune responses prior to patient recovery: a case report of non-severe COVID-19. *Nat Med* 26:453–455. <https://doi.org/10.1038/s41591-020-0819-2>
- UC San Diego Health (2020) <https://health.ucsd.edu/>. Accessed 2020
- Verity R, Okell LC, Dorigatti I, Winskill P, Whittaker C, Imai N, Cuomo-Dannenburg G, Thompson H, Walker PG, Fu H, Dighe A (2020) Estimates of the severity of coronavirus disease 2019: a model-based analysis. *Lancet Infect Dis*. [https://doi.org/10.1016/S1473-3099\(20\)30243-7](https://doi.org/10.1016/S1473-3099(20)30243-7)
- Wang H, Luo S, Shen Y, Li M, Zhang Z, Dong Y, Zhou H, Lin L, Guo W, Kang Z, Xing L (2020) Multiple enzyme release, inflammation storm and hypercoagulability are prominent indicators for disease progression in COVID-19: a multi-centered, correlation study with CT imaging score. *Lancet*. <https://doi.org/10.2139/ssrn.3544837>
- World Health Organization (2020) <https://www.who.int/>. Accessed 2020
- Wu C, Hu X, Song J, Du C, Xu J, Yang D, Chen D, Zhong M, Jiang J, Xiong W, Lang K (2020) Heart injury signs are associated with higher and earlier mortality in coronavirus disease 2019 (COVID-19). medRxiv. <https://doi.org/10.1101/2020.02.26.20028589>
- Yang P, Ding Y, Xu Z, Pu R, Li P, Yan J, Liu J, Meng F, Huang L, Shi L, Jiang T (2020) Epidemiological and clinical features of COVID-19 patients with and without pneumonia in Beijing, China. medRxiv. <https://doi.org/10.1101/2020.02.28.20028068>
- Yun Y, Wang Y, Hao Y, Xu L, Cai Q (2021) The time course of chest CT lung changes in COVID-19 patients from onset to discharge. *Eur J Radiol Open* 8:100305. <https://doi.org/10.1016/j.ejro.2020.100305>
- Zhou F, Yu T, Du R, Fan G, Liu Y, Liu Z, Xiang J, Wang Y, Song B, Gu X, Guan L (2020) Clinical course and risk factors for mortality of adult inpatients with COVID-19 in Wuhan, China: a retrospective cohort study. *Lancet* 395:1054–1062. [https://doi.org/10.1016/S0140-6736\(20\)30566-3](https://doi.org/10.1016/S0140-6736(20)30566-3)

Publisher’s Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.