



# The Promise of AI in Detection, Diagnosis, and Epidemiology for Combating COVID-19: Beyond the Hype

Musa Abdulkareem<sup>1,2,3\*</sup> and Steffen E. Petersen<sup>1,2,3,4</sup>

<sup>1</sup> Barts Heart Centre, Barts Health National Health Service (NHS) Trust, London, United Kingdom, <sup>2</sup> National Institute for Health Research (NIHR) Barts Biomedical Research Centre, William Harvey Research Institute, Queen Mary University of London, London, United Kingdom, <sup>3</sup> Health Data Research UK, London, United Kingdom, <sup>4</sup> The Alan Turing Institute, London, United Kingdom

## OPEN ACCESS

### Edited by:

Jake Y. Chen,  
University of Alabama at Birmingham,  
United States

### Reviewed by:

Maria F. Chan,  
Memorial Sloan Kettering Cancer  
Center, United States  
Akram Mohammed,  
University of Tennessee Health  
Science Center (UTHSC),  
United States

### \*Correspondence:

Musa Abdulkareem  
musa.abdulkareem@nhs.net

### Specialty section:

This article was submitted to  
Medicine and Public Health,  
a section of the journal  
Frontiers in Artificial Intelligence

**Received:** 12 January 2021

**Accepted:** 13 April 2021

**Published:** 14 May 2021

### Citation:

Abdulkareem M and Petersen SE  
(2021) The Promise of AI in Detection,  
Diagnosis, and Epidemiology for  
Combating COVID-19: Beyond the  
Hype. *Front. Artif. Intell.* 4:652669.  
doi: 10.3389/frai.2021.652669

COVID-19 has created enormous suffering, affecting lives, and causing deaths. The ease with which this type of coronavirus can spread has exposed weaknesses of many healthcare systems around the world. Since its emergence, many governments, research communities, commercial enterprises, and other institutions and stakeholders around the world have been fighting in various ways to curb the spread of the disease. Science and technology have helped in the implementation of policies of many governments that are directed toward mitigating the impacts of the pandemic and in diagnosing and providing care for the disease. Recent technological tools, artificial intelligence (AI) tools in particular, have also been explored to track the spread of the coronavirus, identify patients with high mortality risk and diagnose patients for the disease. In this paper, areas where AI techniques are being used in the detection, diagnosis and epidemiological predictions, forecasting and social control for combating COVID-19 are discussed, highlighting areas of successful applications and underscoring issues that need to be addressed to achieve significant progress in battling COVID-19 and future pandemics. Several AI systems have been developed for diagnosing COVID-19 using medical imaging modalities such as chest CT and X-ray images. These AI systems mainly differ in their choices of the algorithms for image segmentation, classification and disease diagnosis. Other AI-based systems have focused on predicting mortality rate, long-term patient hospitalization and patient outcomes for COVID-19. AI has huge potential in the battle against the COVID-19 pandemic but successful practical deployments of these AI-based tools have so far been limited due to challenges such as limited data accessibility, the need for external evaluation of AI models, the lack of awareness of AI experts of the regulatory landscape governing the deployment of AI tools in healthcare, the need for clinicians and other experts to work with AI experts in a multidisciplinary context and the need to address public concerns over data collection, privacy, and protection. Having a dedicated team with expertise in medical data collection, privacy, access and sharing, using federated learning whereby AI scientists hand over training algorithms to the healthcare institutions to train models locally, and taking full advantage of biomedical data stored in biobanks can alleviate some of problems posed by these challenges. Addressing these challenges

will ultimately accelerate the translation of AI research into practical and useful solutions for combating pandemics.

**Keywords:** artificial intelligence, COVID-19, detection, diagnosis, epidemiology, social control, contact tracing, medical imaging

## INTRODUCTION

COVID-19, a type of coronavirus disease caused by Severe Acute Respiratory Syndrome Corona-Virus 2 (SARS-CoV-2), has created enormous suffering, affecting lives and causing deaths. The novel nature of the virus means that humans are only newly exposed to the virus (Brüssow, 2020; Wan et al., 2020). First reported in China in December 2019, it was declared by The World Health Organization (WHO) to be a Public Health Emergency of International Concern (PHEIC) on January 30, 2020 and a pandemic on March 11, 2020 (Team, 2020; WHO, 2020). It is an infectious disease that spreads in humans mainly through respiratory droplets produced by an already infected person through sneezing or talking, or airborne transmission (Moriyama et al., 2020). The early symptoms of the disease include persistent high temperature, dry continuous coughing, loss of taste or smell, and difficulty in breathing (Kooraki et al., 2020; Wang et al., 2020a). Severe cases of the disease cause death (Rothan and Byrareddy, 2020; Zhou et al., 2020a).

Due to the ease with which the coronavirus can spread and grow exponentially within the human population, healthcare resources and manpower to rapidly control it is limited as the number of doctors, nurses, and other healthcare workers and resources that could help control it is finite. Moreover, the disease has exposed weaknesses of many healthcare systems around the world. Indeed, the lack of affordable, quick and accurate means of detecting the disease is one of the most important reasons it has rapidly spread (Ai et al., 2020).

Since the emergence of COVID-19, many governments, research communities, commercial enterprises and other institutions and stakeholders around the world have been fighting in various ways to curb the spread of the disease (Chen et al., 2020a; Dong et al., 2020). Science and technology have helped in the implementation of policies of many governments that are directed toward mitigating the impacts of the pandemic and in developing cures and vaccines for the disease. They also offer unique opportunity to support healthcare workers by providing them with tools that would save them time, improve their ability to carry out their job and enhance the management of healthcare systems developed to combat the pandemic, and much more. Many resources have been made available to support the battle against COVID-19, such as datasets (Cheng et al., 2020; Cohen et al., 2020; Zhao et al., 2020a), computing resource (Hack and Papka, 2020), and research funding (Casigliani et al., 2020; Glasziou et al., 2020; Janiaud et al., 2020; Patel et al., 2020; Prudêncio and Costa, 2020; UKCDR, 2020).

The scope of combating COVID-19 using technology is very broad and it includes understanding the socio-economic and medical impacts of the pandemic. From a healthcare perspective, it includes disease detection, diagnosis, and monitoring (Huang

et al., 2020a; Kong et al., 2020; Thevarajan et al., 2020; Xu et al., 2020a), epidemiology (Chan et al., 2020; Jin et al., 2020a; Li et al., 2020a), social control (Jin et al., 2020a; Kandel et al., 2020; Qian et al., 2020), virology and pathogenesis (Andersen et al., 2020; Jin et al., 2020a; Lu et al., 2020b; Walls et al., 2020), and drug discovery (Chen et al., 2020b; Phua et al., 2020). For example, during the early phase of the outbreak of the pandemic, China used facial recognition cameras to track infected patients and drones to disinfect public places and broadcast audio messages to the public asking them to stay at home (Ruiz Estrada, 2020). As another example, Taiwan linked its national medical insurance database with the immigration and custom database in order to inform the healthcare practitioners of the travel history of patients (Wang et al., 2020b).

The term artificial intelligence (AI) refers to the study of developing computer algorithms with human-like intelligence to accomplish specific tasks. Machine learning (ML) methods are a set of techniques in AI and includes supervised (Kotsiantis et al., 2007), unsupervised (Barlow, 1989), semi-supervised (Zhu, 2005; Chapelle et al., 2009), and reinforcement learning (Sutton and Barto, 1998). Some of these methods and other terms often encountered in the AI literature are briefly described in **Table 1**.

The applications of AI can be found in many disciplines and industries in modern society, and healthcare is not an exemption. The rapid growth of AI-based techniques and tools in healthcare are addressing complex problems such as identifying previously undiscovered relationships in patient phenotypes (Shivade et al., 2014), optimizing healthcare pathways (Lu and Wang, 2019; Blasiak et al., 2020), and improving accuracy of medical decision making (Bennett and Hauser, 2013; Shortliffe and Sepúlveda, 2018).

Advances and accessibility to high-performance scalable computing equipment have driven the recent popularity of the use of AI in many real-world applications. This development has also prompted an expansion of research into novel AI techniques and algorithms. AI algorithms have the potential to interpret biomedical and healthcare data particularly for tasks where conventional statistical methods are less efficient. The algorithms are even more suitable for datasets of large scale and high dimensions. These algorithms can therefore be used to solve problems such as optimizing care pathways, standardizing clinical diagnosis, identifying relationships in patient phenotypes and developing predictive models (Johnson et al., 2017). While AI-based methods can be used to solve many problems in medicine and healthcare, the success of AI projects, in many cases, depends on the choice of the AI technique, the quality of the dataset to be used and the context associated with the way the dataset is used. For instance, deep learning (DL) algorithms such as convolutional neural networks (CNN) are particularly suitable for computer vision problems such as image segmentation (Shen

**TABLE 1** | Some terms and methods commonly used in AI.

	General terms
Artificial Intelligence (AI)	The concept of developing computer algorithms with human-like intelligence to solve specific tasks.
Deep Learning (or Deep Neural Network)	A set of ML algorithms that are based on neural network (NN) that are used for feature learning. The term “deep” refers to the fact that they have multiple layers between the input and the output layers.
Machine Learning (ML)	A subset of AI and consists of a collection of techniques to achieve AI.
Reinforcement Learning	A set of ML algorithms that is based on the interaction between an agent and its environment. In general, the agent seeks to take actions in the environment by maximizing a cumulative reward.
Supervised Learning	A set of ML algorithms for developing mathematical models using data that consists of both the input and the desired output data.
Unsupervised Learning	A set of ML algorithms for finding underlying structures or patterns in datasets using only the input data.
Convolutional Neural Network (CNN)	A set of DL algorithms that are particularly efficient in developing AI-based applications involving images. CNN acts as the backbone of many well-known neural network architectures (such as U-net) used in image processing.
Random Forests (RF) Method	A set of learning algorithms involving several decision trees and whose output is the class that is the statistical mode (in classification tasks) or statistical mean (in regression tasks) of each of the decision trees. These algorithms are often used for classification tasks and regression analysis problems.
Support Vector Machines (SVM)	A set of supervised learning algorithms that constructs hyperplanes in a high-dimensional space. These algorithms are often used for classification tasks, regression analysis, and other problems. In a classification problem, for instance, out of the many hyperplanes, the one that has the largest distance to the data point of any class is considered the ‘optimal’ classifier.

#### Reference List of AI Algorithms Mentioned in this Paper

- AlexNet (Russakovsky et al., 2015; Krizhevsky et al., 2017)
- Artificial Neural Networks (ANN) (Hopfield, 1988; Jain et al., 1996)
- Adaptive-Network-based Fuzzy Inference System (ANFIS) (Jang, 1993)
- CNN (LeCun et al., 2015)
- CNN segmentation model (Region Proposal Network structure) (Ren et al., 2016)
- CNN model with Inception (Szegedy et al., 2016)
- Decision Tree (DT) (Breiman et al., 1984)
- Extreme Gradient Boosting (XGBoost) (Chen and Guestrin, 2016)
- Generative Adversarial Networks (GANs) (Goodfellow et al., 2014)
- Gated Recurrent Unit (GRU) recurrent neural network (Cho et al., 2014; Chung et al., 2014)
- k-mean clustering (Kanungo et al., 2002)
- k-nearest neighbor (Cover and Hart, 1967)
- Least Absolute Shrinkage and Selection Operator (LASSO) logistic regression (Tibshirani, 1996)
- Logistic regression (Hosmer Jr et al., 2013)
- LSTM (Hochreiter and Schmidhuber, 1997)
- RF (Breiman, 2001; Liaw and Wiener, 2002)
- ResNet (He et al., 2016)
- SVM (Cortes and Vapnik, 1995)
- U-Net (Ronneberger et al., 2015)

et al., 2017). Recent advancement in AI research has led to the development of tools in medicine and healthcare that are useful in combating global pandemics. Researchers across several areas of expertise and industries have therefore explored and exploited the use of AI in the battle against COVID-19.

There are many ways in which AI can help in the fight against the COVID-19 pandemic. For example, AI could be used to track the spread of the virus (Al-Qaness et al., 2020; Bandyopadhyay and Dutta, 2020; Carrillo-Larco and Castillo-Cara, 2020; Hu et al., 2020; Jana and Bhaumik, 2020; Huang et al., 2020b; Kavadi

et al., 2020; Sameni, 2020), identify patients with high mortality risk (Jiang et al., 2020a; Qi et al., 2020; Xu et al., 2020b; Yan et al., 2020a), diagnose and screen a population for COVID-19 (Ghoshal and Tucker, 2020; Hassanien et al., 2020; Hemdan et al., 2020; Jin et al., 2020b; Maghdid et al., 2020a; Narin et al., 2020; Wang et al., 2020c,e; Wu et al., 2020a; Zhang et al., 2021; Xu et al., 2020c), or reduce the time for diagnosis (Vaishya et al., 2020a). Many of the AI techniques currently being deployed in the battle already existed prior to the pandemic. These techniques include those that can process and understand medical imaging

data from modalities such as computed tomography (CT) and X-ray that are being used for detection and diagnosis (Li et al., 2020b; Wang et al., 2020e; Wynants et al., 2020) and those involving non-imaging data that are being used for mortality rate and outcome prediction, prognosis, outbreak prediction, contact tracing and social control of COVID-19 (John and Shaiba, 2019; Bandyopadhyay and Dutta, 2020; Chen et al., 2020e; Goh et al., 2020; Pourhomayoun and Shakibi, 2020; Xu et al., 2020b). Other AI techniques have also found new application areas due to COVID-19. For example, in Shi et al. (2020a), argued for the development of AI-based tools for automated acquisition of medical images in order to optimize the imaging workflow and reduce healthcare practitioners' risk of exposure to the virus by minimizing or eliminating contact with COVID-19 patients.

Several reviews, such as Albahri et al. (2020), Bansal et al. (2020), Bragazzi et al. (2020), Bullock et al. (2020), Jamshidi et al. (2020), Kricka et al. (2020), Kumar et al. (2020), Lalmuanawma et al. (2020), Martin et al. (2020), Naudé (2020), Nguyen (2020), Rasheed et al. (2020), Suri et al. (2020), Shi et al. (2020a), Vaishya et al. (2020b), Zhou et al. (2020b), and Chen et al. (2020c), have been published to showcase the opportunities AI presents in the current effort to fight against COVID-19. In this paper, areas where AI techniques are being used in the detection, diagnosis and epidemiological predictions, forecasting and social control for combating COVID-19 are discussed, highlighting areas of successful applications and underscoring issues that need to be addressed to achieve significant progress in battling COVID-19 and future pandemics. The paper assumes a basic background knowledge of AI techniques, the reader is invited to consult (Raghu and Schmidt, 2020) for further information of these AI methods. Useful introduction to the epidemiology and clinical features of COVID-19 can be found in, for example, C Disease Control (2020).

## AI IN COVID-19 DETECTION AND DIAGNOSIS

Detection and diagnosis of COVID-19 is an important part in the fight against the virus. Current diagnostic testing methods are mostly non-invasive methods and they include chest CT and chest X-ray medical imaging, nucleic acid, serologic, and viral throat swab testing methods (Fang et al., 2020; Li et al., 2020c; Lu et al., 2020a; Ozturk et al., 2020; Schwartz, 2020; Zeng et al., 2020). In order to contain the spread of the pandemic and isolate the virus, fast and early detection and tracking of infected patients is crucial and there is clearly the need of innovation in this area (Ai et al., 2020; Fang et al., 2020). In the subsections that follow, AI tools that have been developed for the detection and diagnosis of SARS-CoV-2 and COVID-19 are presented.

### Nucleic Acid Amplification-Based Diagnostics

A type of nucleic acid amplification test (Udugama et al., 2020), the Reverse Transcription-Polymerase Chain Reaction (RT-PCR) test, is one of the most widely used standard testing methods for detecting whether patients have COVID-19 (Ai

et al., 2020). The RT-PCR however suffers from inadequate sensitivity, as low as 71% as reported in Fang et al. (2020), as a result of many factors such as low detection efficiency and complicated sample preparation (Lu et al., 2020a; Wu et al., 2020a). This low sensitivity issue results in multiple testing of many patients usually over several days apart in order to obtain a reliable conclusion.

A ML model was reported in Wu et al. (2020a) that uses 11 key blood indices to distinguish between patients with and without COVID-19. The model was developed using the random forest (RF) ML technique and 49 clinical available blood test parameters (consisting of 24 routine hematological and 25 biochemical parameters) from 169 patients with a total number of 253 data samples of which 105 samples are from patients confirmed to have the COVID-19 disease using the RT-PCR test. The remaining samples consisted of 98 samples from patients with common pneumonia and 25 samples each from patients with tuberculosis and lung cancer. The data was divided into 149 training, 33 testing, and 74 validating datasets. The model achieved accuracy of 96.97% on the testing set and 97.95% for the cross-validation set. While this model (which could be further investigated for reliability and also improved further) offers a promising tool for preliminary assessment of suspected patients with COVID-19, it so far has not made it to front-line in the fight against the coronavirus.

### Medical Imaging Diagnostics

Medical imaging is one of the main areas in which AI has found practical applications in medicine and healthcare. Imaging data obtained using different modalities, such as computed tomography (CT), magnetic resonance imaging (MRI) and X-ray, are of high dimension. They contain very rich information that can be used to develop AI applications. Imaging data can be used to generate many useful image-derived phenotypes that are obtained *via* qualitative and quantitative assessment of structural changes (that often characterize the structural and functional properties of an organ), significantly shortening the time for radiologists to accomplish these tasks (Petersen et al., 2017; Suinesiaputra et al., 2018; Mauger et al., 2019). Imaging data can also be combined with non-imaging data from Electronic Health Record (EHR) or elsewhere for identifying biomarkers and predicting disease risk factors (Alaa et al., 2019). Faster and automated reading and interpretation of image workflow can be achieved using AI-based tools (Petersen et al., 2019; Robinson et al., 2019; Bai et al., 2020).

Furthermore, segmentation of medical images is useful as these images are often affected by noise, artifacts, and other uncertainties associated with imaging. Image segmentation involves contouring a medical image into biologically relevant structures, helping to quantify those structures and their functions and to produce measurements that act as biomarkers (such as quantities that can be used to diagnose, monitor, or prognosticate diseases). In particular, AI-based automatic image segmentation tools are beneficial as they help in eliminating variability that would have been introduced if segmentation were manually done. Consequently, the use of these AI-based technologies has contributed in the fight against COVID-19

(Bullock et al., 2020). Medical imaging modalities, such as chest CT and X-ray imaging, have provided significant support to clinicians in diagnosing COVID-19 (Apostolopoulos and Mpesiana, 2020; Bernheim et al., 2020; Kanne, 2020). A typical workflow for diagnosing COVID-19 with medical imaging modalities involves the following three phases: (i) pre-scan preparation according to a given protocol; (ii) image acquisition; and (iii) diagnosis.

AI tools for COVID-19 diagnosis with medical images often consist of one or a combination of several AI models (or networks) involving the following two main components: (i) image segmentation models, and (ii) image classification. Image segmentation is used to mark and identify the region of interest (ROI) while an image classification task extracts features from the ROI and uses those features as a basis for classifying (diagnosing) the images.

### CT Medical Imaging

Chest CT images are being used for early diagnosis of COVID-19 by identifying ground-glass opacity (GGO) around the subpleural region (Ai et al., 2020; Chung et al., 2020; Fang et al., 2020; Kanne, 2020; Wong et al., 2020). In Pan et al. (2020), the dynamic radiological patterns in chest CT images of COVID-19 patients was reported with the following four stages identified: (i) 0–4 days: early stage; (ii) 5–8 days: progressive stage; (iii) 9–13 days: peak stage; and (iv) 14 days and beyond: absorption stage. These distinct manifestations of COVID-19 in CT images provide evidence and severity of the disease that are exploited using AI systems for diagnosing the disease. Generally, the process of COVID-19 diagnosis with CT images involves the following steps: (i) image pre-processing, (ii) image segmentation, (iii) classification, and (iv) model evaluation.

AI tools for COVID-19 diagnosis with CT images involving lung tissue segmentation are reported in Jin et al. (2020b), Li et al. (2020b), and Xu et al. (2020c). As an example, the AI system presented in Jin et al. (2020b) classifies chest CT input image slices into the following four categories: non-pneumonia, non-viral community acquired pneumonia (CAP), influenza-A/B and COVID-19. The predicted class of an image is that which has the highest probability among the four classes. This AI system was developed using two main DL algorithms: the U-Net for performing the lung segmentation task and the ResNet for performing the classification (diagnosis) task. The ROI in the image include the lung, lung lobes, bronchopulmonary segments, and infected lesions. The dataset consists of 10,250 CT scans from three centers in China and three publicly available external databases. This multi-center dataset was from 7,917 subjects consisting of 3,686 scans of COVID-19, 2,886 scans of CAP, 132 scans of influenza-A/B and 3,546 scans of non-pneumonia subjects. The COVID-19 subjects were all confirmed using the RT-PCR diagnostic test. The imaging dataset of 10,250 was divided into a total training dataset of 5,104 and a total testing dataset of 5,146. As a measure of accuracy, on internal testing dataset of 3,203 images (out of the 5,146) the AI system achieved an AUC of 97.17%, a sensitivity of 90.19% and a specificity of 95.76%. It achieved an AUC of 97.77% on the remaining (external) dataset of 1,943 images.

As another example, the AI system presented in Xu et al. (2020c) classifies chest CT input image slices into the following three categories: influenza-A viral pneumonia (IAVP), COVID-19 and irrelevant to infection (i.e., cases that do not belong to the other two categories) cases. The predicted class of an image is that which has the highest probability that the image belongs to it. This AI system consists of two main DL algorithms: a three-dimensional (3D) CNN segmentation model (Region Proposal Network structure) for performing a lung segmentation task and a ResNet-based model for performing the image classification task. The dataset consists of 618 CT scans from three hospitals in China's Zhejiang Province of which 110 subjects (219 scans) were confirmed of COVID-19 using the RT-PCR diagnostic test; 224 subjects (224 scans) had IAVP, and the remaining 175 scans are healthy subjects. The imaging dataset of 528 (189 COVID-19 cases plus 194 IAVP cases plus 145 healthy cases) were used for training and validation, and the remaining 90 scans (30 COVID-19 cases plus 30 IAVP cases plus 30 healthy cases) were used as a testing dataset. On the testing dataset, the AI system achieved an  $f_1$ -score of 83.9% for COVID-19 cases, 84.7% for IAVP cases, 91.5% for healthy cases and an overall accuracy rate of 86.7%.

Several AI systems, such as Ardakani et al. (2020), Chen et al. (2020d), Gozes et al. (2020), Kang et al. (2020), Li et al. (2020b), Shi et al. (2020b), Song et al. (2020), Tang et al. (2020) and Wang et al. (2020c,d), have been developed for diagnosing COVID-19. Compared to the examples in the two preceding paragraphs, these AI systems mainly differ in their choices of the algorithm for image segmentation of the ROI and the algorithm used for classification or diagnosis. The image segmentation algorithms used include U-Net, U-Net++, V-Net, and others, and the image classification algorithms include ResNet and CNN model with Inception.

In order to address the problem of lack of large datasets of COVID-19 patients for developing AI-based models, researchers, such as in Jin et al. (2020b) and Zhao et al. (2020a), have used different techniques such as data augmentation and transfer learning, to solve the CT image classification problems for COVID-19 diagnosis. In Qian et al. (2020), the classification task was to classify COVID-19 patients into those that will have short-term and long-term hospital stay. Some AI models, such as Shi et al. (2020b,c), went further after the image segmentation task to predict the severity of COVID-19 in patients using algorithms such as least absolute shrinkage and selection operator (LASSO) logistic regression model and RF.

### X-Ray Medical Imaging

The X-ray technology is a very popular imaging modality in medical imaging (Wang et al., 2017). The CT and X-ray medical imaging modalities have been more widely accessible and used to provide evidence and for COVID-19 diagnosis compared to other imaging modalities due to their fast acquisition. In fact, in many healthcare centers and hospitals, X-ray imaging, due to its accessibility and quickness to obtain, is often the first-line imaging modality for suspected COVID-19 patients (Bullock et al., 2020; Shi et al., 2020a). Although the chest X-ray images are less informative compared to CT images for diagnosing COVID-19 due to lower sensitivity of chest X-ray images, the popularity

and availability of X-ray imaging facilities means that it is widely used for the diagnosis of the disease. As with chest CT imaging, chest X-ray imaging is being used for diagnosis of COVID-19 by identifying ground-glass opacity (GGO) around the subpleural region, and these manifestations of COVID-19 in chest X-ray images provide evidence and classification of severity of the disease that are being exploited using AI systems for diagnosing the disease.

In general, the process of COVID-19 diagnosis with chest X-ray images using AI tools involves the following steps: (i) image pre-processing, (ii) image classification, and (iii) model evaluation. In other words, compared to the AI tool for CT images, the image segmentation process is absent although some researchers, such as Hassanien et al. (2020), included classical computer vision methods (i.e., not AI-based methods, such as image thresholding) for carrying out the image segmentation step as well. The AI-based image segmentation part of the process is particularly difficult in the case of chest X-ray images given that the ribs are projected onto other tissues on these images (Chen et al., 2020c) so researchers often skip that step completely. Classification tasks were binary, multi-class, multi-labeled or hierarchical classifications (Albahri et al., 2020).

Several AI systems, such as Ghoshal and Tucker (2020), Hassanien et al. (2020), Hemdan et al. (2020), Maghdid et al. (2020a), Narin et al. (2020), Wang et al. (2019, 2020d), Zhang et al. (2021), have been developed for diagnosing COVID-19 using chest X-ray images. These AI systems mainly differ in their choice of the algorithms used for the image classification task and often combine several algorithms (often, to achieve a feature extraction step before a classification process). The image classification algorithms that are being used include Support Vector Machines (SVM), CNN, AlexNet, ResNet, and CNN model with Inception.

The large number of AI techniques available for diagnosing and classifying a disease means that it can be daunting to select the most appropriate technique (in terms of accuracy and computation efficiency) for a given problem given that many of the researchers have used different (and sometimes conflicting) evaluation criteria for their adopted techniques (Alsalem et al., 2018, 2019; Zaidan et al., 2020). In Albahri et al. (2020), carried out a literature review of AI techniques involving medical images that are being used for diagnosing COVID-19 in an attempt to evaluate and establish benchmarking procedures for these techniques. A detailed description of the proposed methodology for the evaluation and benchmarking of these AI techniques is beyond the scope of this paper and the reader is invited to consult (Albahri et al., 2020) for further information.

## Other Tools for Diagnostics

In Schuller et al. (2020), presented a potential computer audition tool that uses AI-based speech and sound analysis to COVID-19 diagnosis. The authors surveyed automatic recognition and monitoring of contextually significant phenomena from speech or sound, such as dry and wet coughing or sneezing sounds, pain, speech under cold, and breathing for diagnostic exploitation using AI techniques such as Generative Adversarial Networks (GANs) (Pascual et al., 2017).

In Wang et al. (2020f), an AI-based classification model was proposed that is able to distinguish respiratory pattern from six other viral infection respiratory patterns using the Gated Recurrent Unit (GRU) recurrent neural network algorithm with bi-directional attention mechanism. As measures of accuracy, the reported precision, recall,  $f_1$ -score, and accuracy of the model were 94.4, 95.1, 94.8, and 94.5%, respectively. Other models that use respiratory or coughing data for COVID-19 diagnosis can be found in Brown et al. (2020), Imran et al. (2020), and Jiang et al. (2020b).

Researchers, such as in Maghdid et al. (2020b), have also proposed frameworks for using in-built mobile phone sensors including cameras (to scan CT images, for example), temperature sensors, and so on, for COVID-19 diagnosis. The computer audition tools for diagnosing COVID-19, models that use respiratory or coughing data for COVID-19 diagnosis and other AI-based computational frameworks that use speech and sound analysis and in-built mobile sensors, such as Iqbal and Faiz (2020) have not yet gone beyond the conceptual phase.

## AI IN EPIDEMIOLOGY

In the subsections that follow, AI tools that have been developed for epidemiological predictions, forecasting and social control for combating COVID-19 are presented.

### AI for Prognosis

The ability to forecast possible patient outcomes is vital in the planning and management of a pandemic such as COVID-19. In order to improve prognosis and not to overwhelm healthcare systems, the ability to predict number of patients at risk of developing acute respiratory distress syndrome and patients at risk of hospitalization or death can be very important (Bullock et al., 2020). In the fight against MERS Co-V, for example, AI-based models have been used to predict prognosis in patients' infection (in particular, patients' recovery) using patients' profession (e.g., whether healthcare workers or not), age, pre-existing healthcare conditions, and disease severity as model input parameters (John and Shaiba, 2019). Similar AI-based applications and methods have been developed for Ebola patients (Colubri et al., 2016; Riad et al., 2019). These and other similar tools can help, for example, to assess healthcare preparedness for a pandemic and to determine treatment methods and resource allocation during a pandemic, and some of these algorithms could be adapted for decision making in the management of COVID-19 (Bansal et al., 2020).

Epidemiological research is a vast area, and a huge amount of publications on epidemiological modeling of COVID-19 using well-established classical methods have surfaced since the beginning of the pandemic (Cooper et al., 2020; Jewell et al., 2020; Ndairou et al., 2020). Recently, researchers have proposed several AI-based techniques for predicting mortality rate, long-term patient hospitalization (Qi et al., 2020) and patient outcomes for COVID-19 (Jiang et al., 2020a; Yan et al., 2020a). AI-based techniques that have been used to accomplish the prediction tasks include artificial neural networks (ANN), SVM, and XGBoost. For example, in Pourhomayoun and Shakibi

(2020), using dataset of more than 117,000 confirmed COVID-19 patients from 76 countries described in Xu et al. (2020b), the authors used several AI-methods (including SVM, ANN, RF, Decision Tree (DT), logistic regression, and k-nearest neighbor) for the prediction of mortality rate of COVID-19 patients using 112 features consisting of 80 features from patients' doctors notes and health status and 32 features from patients' demographic and physiological data.

## AI for Outbreak Forecasting and Control

The development of forecasting models in order to help policy makers and other stakeholders understand the progression of the pandemic is one of the first areas where mathematical methods were applied to tackle the COVID-19 pandemic. It is therefore not surprising that outbreak forecasting is also one of the first areas in which AI methods have been applied in the fight against the COVID-19 pandemic (Rasheed et al., 2020). There are many existing statistical and dynamic methods for modeling the spread of infectious diseases and understand the impact of interventions to curb these diseases, such as mass vaccination or social distancing, in any given population (Anderson and May, 1979; May and Anderson, 1979; Mena-Lorcat and Hethcote, 1992; Isham and Medley, 1996; Vynnycky and White, 2010; Siettos and Russo, 2013; Pastor-Satorras et al., 2015). Several of these methods have been used to understand and forecast the spread of COVID-19 from available data (Karako et al., 2020; Sameni, 2020; Wu et al., 2020b; Zhao et al., 2020b). These methods can be used to determine transmission factors in order to establish preventive and control measures for the pandemic.

The majority of AI applications developed in the fight against COVID-19 have focused on predicting national and local statistics such as the number of confirmed cases, deaths, and people recovered from COVID-19 (Bullock et al., 2020). AI models that have been developed for outbreak predictions include (Al-Qaness et al., 2020; Bandyopadhyay and Dutta, 2020; Carrillo-Larco and Castillo-Cara, 2020; Hu et al., 2020; Jana and Bhaumik, 2020; Huang et al., 2020b; Kavadi et al., 2020; Sameni, 2020), and the modeling techniques used for these models include CNN, long short-term memory (LSTM), adaptive-network-based fuzzy inference system (ANFIS), partial derivative regression and non-linear machine learning (PDR-NML) (Kavadi et al., 2020), SVM and k-mean clustering.

For example, in Carrillo-Larco and Castillo-Cara (2020), a model based on the k-means clustering algorithm was developed and used to categorize countries based on the number of confirmed COVID-19 cases using a dataset that contains features such as the prevalence of HIV/AIDS, diabetes, and tuberculosis in 156 countries in addition to data on the number of COVID-19 related deaths, confirmed cases and recovered cases. In Al-Qaness et al. (2020), an ANFIS-based model was developed to estimate and forecast the number of confirmed cases of COVID-19 10 days ahead using data of previously confirmed cases. And in Ribeiro et al. (2020), for 10 Brazilian states with a high daily COVID-19 incidence, a stacked ensemble of learning algorithms [autoregressive integrated moving average (ARIMA), cubist regression (CUBIST), RF, ridge regression (RIDGE), SVM] with a Gaussian process (GP) meta-learner was used to conduct

1, 3, and 6-days ahead time series forecasting of the COVID-19 cumulative confirmed cases, achieving errors in a range of 0.87–3.51%, 1.02–5.63%, and 0.95–6.90% in 1, 3, and 6-days-ahead predictions, respectively.

In addition, some of these AI-based models, such as in Kavadi et al. (2020), have reported prediction accuracies that are superior to traditional linear regression-based methods. Researchers, such as in Fong et al. (2020), have also proposed techniques for comparing these different models that have mostly been developed using different architectures and trained with non-identical hyperparameters.

## AI for Contact Tracing and Social Control

The implementation of indiscriminate lockdowns in several countries in an attempt to control the COVID-19 pandemic have had severe social and economic consequences. Despite the physical distancing measures in-place when some of the lockdown restrictions were gradually relaxed, other public health measures were necessary in order to control the pandemic (Hellewell et al., 2020; Hope et al., 2020; Park et al., 2020; Salathé et al., 2020; Kretzschmar et al., 2020a,b), and contact tracing (whether conventional methods that rely on interviewing COVID-19 patients or mobile phone application technology) has been one of the methods that have been adopted in many parts of the world for this purpose. Contact tracing involves contact identification, contact listing and contact follow-up (Kricka et al., 2020).

For contact tracing purposes, mobile applications that have been deployed to notify every participating user that a person with COVID-19 was within a certain distance of the user for more than a specific amount of time include COVIDSafe (Australia), Ketju (Finland), CoronaApp (Germany), StopCovid (France), NZ COVID Tracer (New Zealand), TraceTogether (Singapore), NHS Covid-19 App (United Kingdom), to mention a few (Lalmuanawma et al., 2020). As far as we know, none of these digital technologies have been confirmed to use AI-based models as tools, for example, in identifying those in contact with a COVID-19 patient [in Lalmuanawma et al. (2020) though, there is a report that AI tools are being used but this could not be confirmed in the references provided by the authors]. There are however promises [see Kricka et al. (2020), for example] that data gathered through these applications could be exploited for developing AI-based tools in the future.

In addition, AI techniques have been used to develop applications for managing and control the spread of the COVID-19 pandemic. Technologies, such as drones and surveillance cameras equipped with AI-based models for enforcing social isolation (Ahmed et al., 2021), have been reported. As for the impacts of the various social control strategies, the reader is invited to consult (Chang et al., 2020; Hellewell et al., 2020; Kissler et al., 2020; Koo et al., 2020) for further information.

## DISCUSSION

Promising and encouraging AI-based techniques and frameworks for the detection, diagnosis, and epidemiological predictions, forecasting and social control of COVID-19 have

been proposed in the fight against the disease. For these AI techniques to gain wide acceptance and use in practical clinical settings however, there would need to be a framework on how these models would be incorporated into clinical practice systems. Importantly, the models, which have been developed with mostly limited amount of data using different algorithms and architectures, would need to be trained and validated with larger amount of data and issues such as overfitting and biasness should be appropriately addressed. Evaluating and comparing the performance of AI models is crucial but challenging. This is partly due to complex relationships amongst the choice of algorithms, architectures, hyperparameters, and the quality and amount of data used for these models.

In addition, many (if not the majority) of the proposed or developed AI-based techniques and models for COVID-19 diagnosis and epidemiological forecasting have not been externally evaluated and caution must be exercised in the interpretation of these results. Indeed, despite the urgency for the publication of research results during the COVID-19 pandemic, these models cannot be used in clinical practice in their current form as critical review and external assessment of the techniques and models with multi-center datasets should be carried out.

To illustrate the scale of the lack of external evaluation problem with an example, consider a recent study presented in Yan et al. (2020b) where the authors have used blood samples from 485 infected patients in the region of Wuhan, China, to identify crucial predictive biomarkers of disease mortality using AI-based tools. In this relatively simple severity and outcome prediction task, and with a small validation sample size and no external model evaluation, the authors have used the XGBoost classifier method to identify three biomarkers [namely, lactic dehydrogenase (LDH), lymphocyte count and high-sensitivity C-reactive protein (hs-CRP)] that will allow the prediction of the mortality of COVID-19 patients more than 10 days in advance with reportedly more than 90% accuracy. External evaluation of this result by several other researchers, such as in Barish et al. (2020), Giacobbe (2020), Quanjel et al. (2020), and Dupuis et al. (2021), has shown that the results of Yan et al. (2020b) have limited clinical utility as it was impossible to replicate the findings and arrive at the same conclusion. If a huge external evaluation problem exists even for simpler problems (such as prediction and forecasting problems), one can only imagine the scale of the problem when using AI-based model for more complicated problems such as those involving images (computer vision-related problems).

AI has huge potential in the battle against the COVID-19 pandemic. Despite several AI approaches and techniques proposed for the detection, diagnosis and epidemiological predictions, forecasting, and social control in the combat against the pandemic, successful practical deployments of these AI-based tools have so far been limited. There are challenges that have led to the limited applicability of these AI-based tools. In the following paragraphs, some of these challenges are discussed with some suggestions of how some of these obstacles may be tackled in order to achieve significant progress in battling COVID-19 and future pandemics using AI techniques.

## Data Accessibility

One of the key challenges that AI experts have faced during the COVID-19 pandemic is the lack of access to sufficiently large datasets for training and external validation of AI models upon which deployable and successful applications depend. In order to tackle this problem for COVID-19 and future pandemics, healthcare centers would need a dedicated team with expertise in medical data collection, privacy, access, and sharing. In short, data governance frameworks and protocols for pandemics and other emergency times will need to be designed and put in place.

One of the sources of data that has not been taken full advantage of so far for developing AI-based applications and solutions during the COVID-19 pandemic are data from biobanks. Biobanks provide infrastructure for the collection and storage of biomedical data, including data related to health records and lifestyle of participants, with the aim of advancing scientific research and improving healthcare. They are often large databases that can store imaging data, text data from electronic health record (EHR) and lifestyle information, and numerical data obtained by physical measurements of consented participants. Several types of biobanks exist around the world with different population sizes, including genetic banks, blood banks, and tissue banks. These biobanks contain valuable data that can provide insights into how the health of a population develops over years and provide a rich source of data that can be harnessed to unveil complex relationships amongst variables [such as environmental (Wright et al., 2002; Hall et al., 2014), lifestyle choice (Rutten-Jacobs et al., 2018; Said et al., 2018), and genetics (Arnau-Soler et al., 2019; Wang et al., 2019)] that are associated with COVID-19. Biobanking is particularly useful in that it provides a unified data repository with mostly standardized data collecting protocols. In contrast, the hospital data are “messy” due to the nature of data collection and storage across multiple repositories. Examples of biobanks include the Kaiser Permanente’s Research Program on Genes, Environment and Health (RPGEH) with 200,000 participants (Kaiser Permanente, 2020), the UK Biobank with 500,000 participants (Biobank, 2014), China Kadoorie Biobank with 500,000 participants (Chen et al., 2005a, 2011), India’s Chennai biobank with 500,000 participants (Gajalakshmi et al., 2007), and Biobank of Vanderbilt University Medical Center (BioVU) with over 1.4 million participants (Roden et al., 2008). Not all these biobanks have data of COVID-19 patients. The UK Biobank, one of the largest biobanks in the world in terms of data volume and depth including multi-organ imaging, is an example of one that has been integrated with pre-existing data of COVID-19 patients. UK Biobank’s data has been used for research related to COVID-19 [for example, see Armstrong et al. (2020), Atkins et al. (2020), Grant and McDonnell (2020), Hastie et al. (2020), Jimenez-Solem et al. (2020), Kenneth and So (2020), Pereira et al. (2020), Sattar et al. (2020), Toh and Brody (2020), and Zimmerman and Kalra (2020)]. Few AI-based applications, such as in Jimenez-Solem et al. (2020), Kenneth and So (2020), Pereira et al. (2020), Toh and Brody (2020), and Zimmerman and Kalra (2020), exist that have used biobanks’ data for their development, and it is likely

that the use of biobanks' data for the development of AI solutions will increase in the near future.

In addition, researchers, such as in Brisimi et al. (2018), Lee et al. (2018), Rieke et al. (2020), Li et al. (2020d), and Xu et al. (2020d), have proposed the use of federated learning (FL) whereby, rather than participating healthcare institutions hand over healthcare data to AI experts to develop AI models, AI experts will handover training algorithms to the healthcare institutions to train their models locally. The AI experts only get the model or model parameters in return—thus, eliminating some of the problems of data governance and privacy associated with data transfer between different parties while giving access to large amount of data. FL is not without its challenges, such as lesser accuracy of the final model (Li et al., 2020d), and the reader is invited to consult (Brisimi et al., 2018; Lee et al., 2018; Li et al., 2020d; Xu et al., 2020d) for further information of this approach.

## External Evaluation

Many of the developed AI-based techniques and models for COVID-19 diagnosis and epidemiological forecasting have not been externally evaluated. External model evaluation helps in assessing the generalisability of the predictions on independent datasets and ensures that the model has learnt the underlying features of the process that produces the data rather than “memorized” the features of a particular set of data. For illustration, **Figure 1** shows the steps in developing models using AI algorithms, highlighting the model evaluation stage of the development process.

Many publicly available datasets for COVID-19 diagnosis do not necessarily generalize to the whole population (i.e., they are usually for a specific country or regions of a country or a specific number of hospitals). The implication is that most ML models based on them will be biased (He and Garcia, 2009), which can reduce the performance of the models in practical settings (Chawla et al., 2002) and can promote healthcare inequalities (Petersen et al., 2019). Many mainstream ML algorithms for classification problems, including SVM, decision trees, and nearest neighbor, were developed based on the assumption that the dataset has balanced class distribution (Chen et al., 2005b; Almogahed and Kakadiaris, 2015), resulting in significant error when classifying the minority class. Algorithms that have been developed to overcome this problem algorithmically or at data level can be found in Hart (1968), Kubat and Matwin (1997), Laurikkala (2001), Barandela et al. (2003), Oh (2011), and Almogahed and Kakadiaris (2015). In addition, it is important for published research to report the pre-processing, the cleaning and the feature engineering steps applied to the data used for developing AI-based solutions.

## AI Regulatory Landscape

Recently, frameworks for strong regulatory and ethical requirements of AI-based clinical utility tools are being developed but significant hurdles still persist (Petersen et al., 2019). Many AI experts are unaware of the regulatory landscape governing the development of AI tools in healthcare and have not considered this matter in their development. Proof of model performance is not sufficient. Issues, such as model biasness,

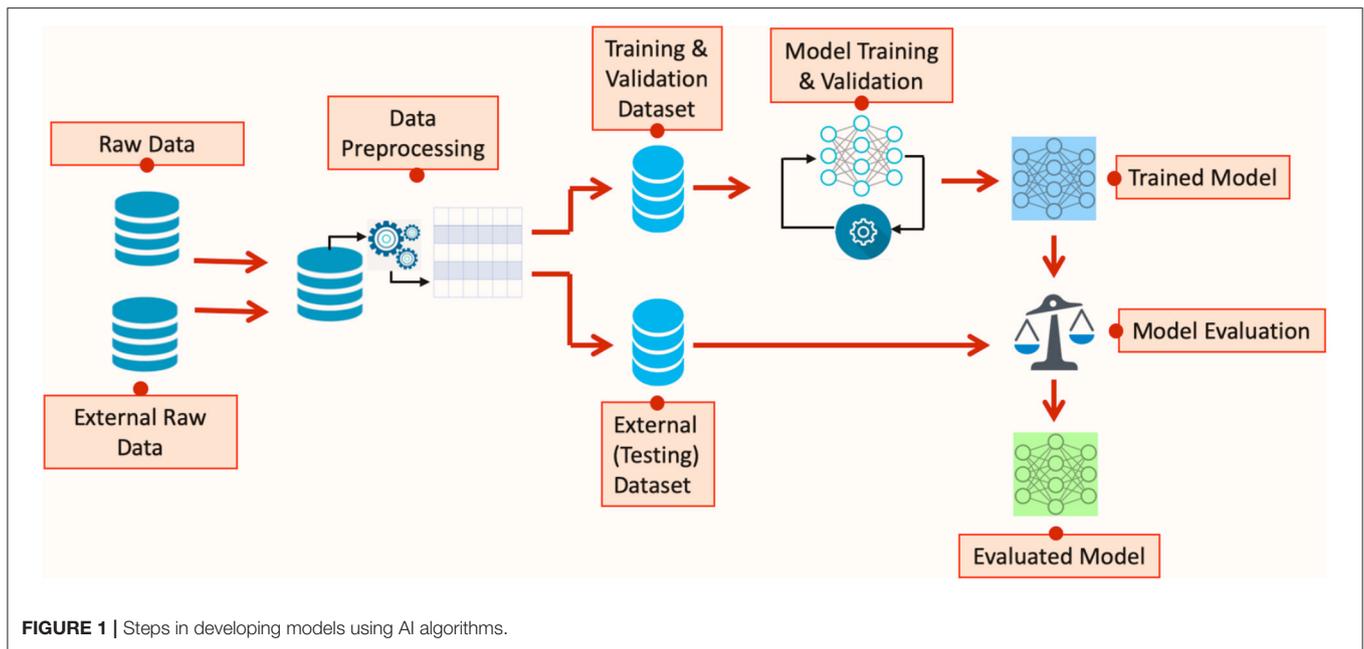
safety, effectiveness, and benefit-versus-harm analysis have mostly been ignored by developers of many critical AI-based healthcare technologies.

## Collaboration Between AI Experts and Clinicians

While AI provides the opportunity to reduce the time for disease diagnosis and improve accuracy, the workload of healthcare professionals is very high during a pandemic. The impact of this includes the difficulty for healthcare professionals to be up to date with the progress being made in areas relevant to their work. It has also hindered their contribution toward that progress. The absence and lack of engagement of clinicians to contribute and review research results during the COVID-19 pandemic has contributed to the limited impact, reliability and clinical utility of many of these research findings. The COVID-19 pandemic has highlighted the importance of domain specific knowledge in AI. It is not sufficient for clinicians to handover data to AI experts who understand how to develop and use classical AI algorithms. Rather, it is important for the clinicians to work with AI experts to help them understand the context of the solutions being developed, to help them interpret the results from those solutions, and to guide them on how those solutions could be used and integrated into existing clinical healthcare pathways or workflows. Thus, multidisciplinary research collaborations will no doubt accelerate the translation of AI research into practical solutions in healthcare and funding bodies could help in this by ensuring multidisciplinary collaborations as a condition for funding. An important lesson that should be learnt in using AI techniques for combating COVID-19 and future pandemics is that the applicability of these techniques is limited if AI experts work in isolation. Important progress in healthcare using AI technologies can be achieved only in a multi-disciplinary setting where clinicians, epidemiologists, computer scientists, software developers, AI experts and others work together to achieve the common goal of improving healthcare services through innovative technologies.

## Public Engagements Over Privacy Concerns

While AI-based technologies embedded in digital systems have played a role in controlling the spread of COVID-19 and the general management of the disease by many governments across the world, the concerns of the general public over privacy have had an impact on the acceptance of many of these technologies and even other potential applications. Consider, for instance, the contract tracing mobile applications that many governments have deployed as a tool for controlling COVID-19, concerns over the possibility that data gathered through these applications could be exploited for other purposes has meant that the general public have been very reluctant in using them (Clark et al., 2020; Lewis, 2020). It has also meant that tools applicable to one country (such as China's use of facial recognition cameras to track infected patients or the linking of the national medical insurance database with the immigration and custom database in order to



**FIGURE 1** | Steps in developing models using AI algorithms.

inform the healthcare practitioners in Taiwan of the travel history of patients) may not be applicable to others.

A framework that will ensure transparency over the legal basis of data use, that data collection is safe and that there are controls and mechanisms to protect misuse of data is critical now and in future. Thus, while it is essential to gather data to address the challenges posed by a pandemic, the authorities would need to do work on gaining the trust of the population through effective engagements with all stakeholders on the mechanisms that would be in place in order to protect privacy and data misuse.

## Potential Misuse of AI Applications

One of the dangers of reliance on AI applications during a pandemic is the potential for misuse. Medical imaging involves several stages including image acquisition, reconstruction, and transmission for storage using Digital Imaging and Communications in Medicine (DICOM) protocol. A cyber-attack could disrupt the use of the devices such as CT devices that can be critical for disease diagnosis during a pandemic (Mahler et al., 2018). With the advent of advanced AI techniques such as generative adversarial network in medical imaging (Yi et al., 2019), one can envisage sophisticated scenarios where AI technologies are used for cyber-attacks that can alter the output of imaging modalities (for instance, by removing or adding a tissue to medical images) altering the results of medical examination, which could lead to fatal consequences. With increasing cyber-attack activities during COVID-19 (Lallie et al., 2021; Muthuppalaniappan and Stevenson, 2021), healthcare providers must be prepared for preventing the occurrence and also detecting and mitigating the impacts that these type of AI attacks will cause when they occur.

In addition, while FL can resolve data governance issues, it does not necessarily guarantee data security on its own as it may be possible to reconstruct parts of the training dataset

from the weights on decentralized computer nodes (Kaissis et al., 2020). This possibility can allow attackers to steal sensitive personal information in the training datasets from the nodes and even reconstruct medical images with high degree of accuracy (Fredrikson et al., 2015; Hitaj et al., 2017), leading to patient confidentiality violations.

Problems associated with data imbalance, variability and incompleteness resulting from the use of datasets that are not accurate representation of the population on which AI models was built for can lead to biased treatment of certain ethnic, sex, age, and other groups. In many cases, these data biases are often introduced inadvertently by AI algorithm developers but unscrupulous individuals can take advantage of this to exacerbate bias from cultural prejudices and increase disparities in delivering healthcare services. Moreover, misuse of AI models can also result when the datasets used for model training do not take into account future use-case conditions; for example, radiologists can easily adapt to change in MRI field strength and breathing motion artifacts but these changes will affect the performance of AI models unless they have been specifically allowed for during the training of the models (Brady and Neri, 2020).

These issues of misuse of AI as highlighted here show that it is important to provide safeguards to ensure that new AI solutions during a pandemic are assessed before being deployed at scale. It is important to emphasize that these challenges posed by AI are not necessarily associated with the limitations of AI *per se* (Rodriguez et al., 2018). Rather, they apply to particular use-cases and emphasize the importance of understanding the relationships that AI models use in arriving at their predictions. As such, guards against spurious predictions must be put in place in order to limit data misuse.

We finish by noting that, recently, there have been several promising initiatives from key players (e.g., government bodies,

commercial institutions, and policy makers) to collect and manage data in order to address or alleviate some of the problems highlighted in this paper. We mention a few of them in the following paragraphs.

In the United Kingdom, NHSX, the government's unit responsible for developing and setting national policy on digital, data and technology for National Health Service (NHS), has developed the National COVID-19 Chest Imaging Database (NCCID) in order to collect patient data and facilitate research and the development and validation of technologies that are promising for improving COVID-19 care (NHSX, 2021). The categories of collected data include chest X-ray, CT, and MR images including those performed in the 3 years preceding the first COVID-19-related imaging study, routine demographic data, biochemical and hematological data, and outcome data.

In the European Union (EU), SoBigData is a research initiative under the EU's Horizon 2020 programme (Grant No. 654024 and 871042) which provides an integrated ecosystem of "big data" for ethnic-sensitive scientific discoveries in multiple fields including mathematics, ICT, and human, social, and economic sciences (SoBigData.eu, 2021). The idea is to promote repeatable and open science by meeting the data and infrastructural needs of researchers while also ensuring that users' data are gathered for specific application and timebound (e.g., relates to dealing with the COVID-19 pandemic only and data will be deleted afterwards), the data cannot be shared without consent, and the data must be of direct benefit to the users whose data were gathered. Another initiative in the EU, The Confederation of Laboratories for Artificial Intelligence (CLAIRE) (CLAIRE, 2021), has warned that it is very likely that our societies will be confronted with other crises at a scale similar to COVID-19 in the not-so-far future and have outlined an European approach with the recommendation that standards and frameworks that would facilitate the development of efficient management of medical data that will not erode human dignity must be developed (Ishmaev et al., 2021).

In the United States, three national institutions namely, National Center for Advancing Translational Sciences (NCATS), Clinical and Translational Science Awards (CTSA) Program and Institutional Development Award Networks for Clinical and Translational Research (IDeA-CTR), have partnered to form the National COVID Cohort Collaborative (N3C) in an attempt to enable collaborators to contribute and use COVID-19 clinical data for scientific research that will have impact in the battle against the pandemic (NCATS-US, 2021). As at the time of writing this paper, the data of more than 950,000 COVID-19 positive patients are available from N3C for researchers to examine associations between COVID-19 patient outcomes and other determinants of health and, at least, 144 projects are already on-going for this purpose. Interestingly, in addition to patient data being de-identified for privacy reasons, this cloud-based data repository consists of synthetic (that is, computationally derived) data that statistically resemble original patient information but are not the actual data of the patients, adding another layer of privacy protection for patients.

The summary of the key messages and the main lessons learnt on the application of AI-based techniques and frameworks

for the detection, diagnosis and epidemiological predictions, forecasting and social control of COVID-19 is as follows:

- We recommend that healthcare centers set up dedicated teams with expertise in medical data collection, privacy, access and sharing, and data governance frameworks and protocols for pandemics and other emergency times.
- External model evaluation is important to avoid the problems associated with model overfitting and biasness, such as arriving at clinically unusable solutions or introducing inequalities in health and healthcare. We recommend the establishment of independent units at national level or through international collaboration with the goal of assessing and validating AI applications developed for healthcare during pandemics before such applications are adopted and scaled up.
- The regulatory landscape (covering issues such as safety, effectiveness and benefit-versus-harm analysis) governing the development of AI tools in healthcare need to be accessible and understandable to AI experts. We recommend that professional bodies that will oversee certification programmes for AI experts working in healthcare be introduced to ensure that, through continuing professional development, these professionals adhere to common ethical standards and are aware of the current ethical and social issues related to their work.
- The COVID-19 pandemic has highlighted the importance of domain specific knowledge in AI, and multidisciplinary research collaborations will only accelerate the translation of AI research into practical and useful solutions in healthcare. In funding AI projects, we recommend that research fund awarding bodies should make the collaboration between AI scientists and domain specific experts a condition for grant awards.
- In order to gain the trust of the population in terms of data collection, privacy and protection, we recommend that all stakeholders work together in the development of a data use and sharing framework that will ensure effective data management is in place for the development and advancement of AI applications in healthcare.

## CONCLUSION

In this paper, AI techniques that are being used in the detection, diagnosis and epidemiological predictions, forecasting and social control for combating COVID-19 have been discussed. While AI has huge potential in the battle against COVID-19, the successful practical deployments of these AI-based tools have so far been limited due to challenges such as limited data accessibility, need for external evaluation of AI models, lack of awareness of AI experts of the regulatory landscape governing the deployment of AI tools in healthcare, the need for clinicians and other experts to work with AI experts in a multidisciplinary context and the need to address public concerns over data collection, privacy and protection. Overcoming these challenges will lead to significant progress in battling COVID-19 and future pandemics using AI techniques.

## AUTHOR CONTRIBUTIONS

MA drafted the first version of the manuscript. MA and SP contributed to the content and writing of the final version. Both authors contributed to the article and approved the submitted version.

## FUNDING

MA and SP acknowledge support from the CAP-AI programme (led by Capital Enterprise in partnership with Barts Health

NHS Trust and Digital Catapult and funded by the European Regional Development Fund and Barts Charity) and Health Data Research UK (HDR UK—an initiative funded by UK Research and Innovation, Department of Health and Social Care (England) and the devolved administrations, and leading medical research charities; [www.hdruk.ac.uk](http://www.hdruk.ac.uk)). SP acknowledges support from the National Institute for Health Research (NIHR) Biomedical Research Center at Barts and from the SmartHeart EPSRC programme grant ([www.nihr.ac.uk](http://www.nihr.ac.uk); EP/P001009/1). SP has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No. 825903 (euCanSHare project).

## REFERENCES

- Ahmed, I., Ahmad, M., Rodrigues, J., Jeon, G., and Din, S. (2021). A deep learning-based social distance monitoring framework for COVID-19. *Sustain. Cities Soc.* 65:102571. doi: 10.1016/j.scs.2020.102571
- Ai, T., Yang, Z., Hou, H., Zhan, C., Chen, C., Lv, W., et al. (2020). Correlation of chest CT and RT-PCR testing in coronavirus disease 2019 (COVID-19) in China: a report of 1,014 cases. *Radiology* 2020:200642. doi: 10.1148/radiol.2020200642
- Alaa, A. M., Bolton, T., Di Angelantonio, E., Rudd, J. H. F., and van der Schaar, M. (2019). Cardiovascular disease risk prediction using automated machine learning: a prospective study of 423,604 UK Biobank participants. *PLoS ONE* 14:e0213653. doi: 10.1371/journal.pone.0213653
- Albahri, O. S., Zaidan, A. A., Albahri, A. S., Zaidan, B. B., Abdulkareem, K. H., Al-Qaysi, Z. T., et al. (2020). Systematic review of artificial intelligence techniques in the detection and classification of COVID-19 medical images in terms of evaluation and benchmarking: taxonomy analysis, challenges, future solutions and methodological aspects. *J. Infect. Public Health* 13, 1381–1396. doi: 10.1016/j.jiph.2020.06.028
- Almogahed, B. A., and Kakadiaris, I. A. (2015). NEATER: filtering of over-sampled data using non-cooperative game theory. *Soft Comput.* 19, 3301–3322. doi: 10.1007/s00500-014-1484-5
- Al-Qaness, M. A. A., Ewees, A. A., Fan, H., and Abd El Aziz, M. (2020). Optimization method for forecasting confirmed cases of COVID-19 in China. *J. Clin. Med.* 9:E674. doi: 10.3390/jcm9030674
- Alsalem, M. A., Zaidan, A. A., Zaidan, B. B., Albahri, O. S., Alamoodi, A. H., Albahri, A. S., et al. (2019). Multiclass benchmarking framework for automated acute Leukaemia detection and classification based on BWM and group-VIKOR. *J. Med. Syst.* 43:212. doi: 10.1007/s10916-019-1338-x
- Alsalem, M. A., Zaidan, A. A., Zaidan, B. B., Hashim, M., Albahri, O. S., Albahri, A. S., et al. (2018). Systematic review of an automated multiclass detection and classification system for acute Leukaemia in terms of evaluation and benchmarking, open challenges, issues and methodological aspects. *J. Med. Syst.* 42:204. doi: 10.1007/s10916-018-1064-9
- Andersen, K. G., Rambaut, A., Lipkin, W. I., Holmes, E. C., and Garry, R. F. (2020). The proximal origin of SARS-CoV-2. *Nat. Med.* 26, 450–452. doi: 10.1038/s41591-020-0820-9
- Anderson, R. M., and May, R. M. (1979). Population biology of infectious diseases: Part I. *Nature* 280, 361–367. doi: 10.1038/280361a0
- Apostolopoulos, I. D., and Mpesiana, T. A. (2020). Covid-19: automatic detection from x-ray images utilizing transfer learning with convolutional neural networks. *Phys. Eng. Sci. Med.* 2020:1. doi: 10.1007/s13246-020-00865-4
- Ardakani, A. A., Kanafi, A. R., Acharya, U. R., Khadem, N., and Mohammadi, A. (2020). Application of deep learning technique to manage COVID-19 in routine clinical practice using CT images: results of 10 convolutional neural networks. *Comput. Biol. Med.* 2020:103795. doi: 10.1016/j.compbiomed.2020.103795
- Armstrong, J., Rudkin, J. K., Allen, N., Crook, D. W., Wilson, D. J., Wyllie, D. H., et al. (2020). Dynamic linkage of covid-19 test results between public health england's second generation surveillance system and uk biobank. *Microb. Genom.* 6:397. doi: 10.1099/mgen.0.000397
- Arnau-Soler, A., Macdonald-Dunlop, E., Adams, M. J., Clarke, T.-K., MacIntyre, D. J., Milburn, K., et al. (2019). Genome-wide by environment interaction studies of depressive symptoms and psychosocial stress in UK Biobank and Generation Scotland. *Transl. Psychiatry* 9, 1–13. doi: 10.1038/s41398-018-0360-y
- Atkins, J. L., Masoli, J. A. H., Delgado, J., Pilling, L. C., Kuo, C.-L., Kuchel, G. A., et al. (2020). Preexisting comorbidities predicting COVID-19 and mortality in the UK Biobank Community Cohort. *J. Gerontol. Ser. A* 75, 2224–2230. doi: 10.1093/gerona/glaa183
- Bai, W., Suzuki, H., Huang, J., Francis, C., Wang, S., Tarroni, G., et al. (2020). A population-based phenome-wide association study of cardiac and aortic structure and function. *Nat. Med.* 2020, 1–9. doi: 10.1038/s41591-020-1009-y
- Bandyopadhyay, S. K., and Dutta, S. (2020). Machine learning approach for confirmation of covid-19 cases: positive, negative, death and release. *MedRxiv*. doi: 10.2196/preprints.19526
- Bansal, A., Padappayil, R. P., Garg, C., Singal, A., Gupta, M., and Klein, A. (2020). Utility of artificial intelligence amidst the COVID 19 pandemic: a review. *J. Med. Syst.* 44, 1–6. doi: 10.1007/s10916-020-01617-3
- Barandela, R., Sánchez, J. S., García, V., and Rangel, E. (2003). Strategies for learning in class imbalance problems. *Pattern Recognit.* 36, 849–851. doi: 10.1016/S0031-3203(02)00257-1
- Barish, M., Bolourani, S., Lau, L. F., Shah, S., and Zanos, T. P. (2020). External validation demonstrates limited clinical utility of the interpretable mortality prediction model for patients with COVID-19. *Nat. Mach. Intell.* 2020, 1–3. doi: 10.1038/s42256-020-00254-2
- Barlow, H. B. (1989). Unsupervised learning. *Neural Comput.* 1, 295–311. doi: 10.1162/neco.1989.1.3.295
- Bennett, C. C., and Hauser, K. (2013). Artificial intelligence framework for simulating clinical decision-making: a Markov decision process approach. *Artif. Intell. Med.* 57, 9–19. doi: 10.1016/j.artmed.2012.12.003
- Bernheim, A., Mei, X., Huang, M., Yang, Y., Fayad, Z. A., Zhang, N., et al. (2020). Chest CT findings in coronavirus disease-19 (COVID-19): relationship to duration of infection. *Radiology* 2020:200463. doi: 10.1148/radiol.2020200463
- Biobank, U. K. (2014). *About UK Biobank*. Available online at: <https://www.ukbiobank.ac.uk/learn-more-about-uk-biobank/about-us> (accessed January 12, 2021).
- Blasiak, A., Khong, J., and Kee, T. (2020). CURATE. AI: optimizing personalized medicine with artificial intelligence. *SLAS Technol. Transl. Life Sci. Innov.* 25, 95–105. doi: 10.1177/2472630319890316
- Brady, A. P., and Neri, E. (2020). Artificial intelligence in radiology—ethical considerations. *Diagnostics* 10:231. doi: 10.3390/diagnostics10040231
- Bragazzi, N. L., Dai, H., Damiani, G., Behzadifard, M., Martini, M., and Wu, J. (2020). How big data and artificial intelligence can help better manage the COVID-19 pandemic. *Int. J. Environ. Res. Public Health* 17:3176. doi: 10.3390/ijerph17093176
- Breiman, L. (2001). Random forests. *Mach Learn* 45, 5–32. doi: 10.1023/A:1010933404324
- Breiman, L., Friedman, J., Stone, C. J., and Olshen, R. A. (1984). *Classification and Regression Trees*. Boca Raton, FL: CRC Press.

- Brisimi, T. S., Chen, R., Mela, T., Olshevsky, A., Paschalidis, I. C., and Shi, W. (2018). Federated learning of predictive models from federated electronic health records. *Int. J. Med. Inform.* 112, 59–67. doi: 10.1016/j.ijmedinf.2018.01.007
- Brown, C., Chauhan, J., Grammenos, A., Han, J., Hasthanasombat, A., Spathis, D., et al. (2020). Exploring automatic diagnosis of COVID-19 from crowdsourced respiratory sound data. *ArXiv Prepr ArXiv200605919*. doi: 10.1145/3394486.3412865
- Brússow, H. (2020). The novel coronavirus – a snapshot of current knowledge. *Microb. Biotechnol.* 13, 607–612. doi: 10.1111/1751-7915.13557
- Bullock, J., Pham, K. H., Lam, C. S. N., Luengo-Oroz, M. (2020). Mapping the landscape of artificial intelligence applications against COVID-19. *ArXiv Prepr ArXiv200311336*. doi: 10.1613/jair.1.12162
- C Disease Control (2020). *Coronavirus Disease 2019 (COVID-19): Frequently Asked Questions*. Webpage.
- Carrillo-Larco, R. M., and Castillo-Cara, M. (2020). Using country-level variables to classify countries according to the number of confirmed COVID-19 cases: an unsupervised machine learning approach. *Wellcome Open Res.* 5:56. doi: 10.12688/wellcomeopenres.15819.1
- Casigliani, V., De Nard, F., De Vita, E., Arzilli, G., Grosso, F. M., Quattrone, F., et al. (2020). Too much information, too little evidence: is waste in research fuelling the covid-19 infodemic. *BMJ* 370:m2672. doi: 10.1136/bmj.m2672
- Chan, J. F.-W., Yuan, S., Kok, K.-H., To, K. K.-W., Chu, H., Yang, J., et al. (2020). A familial cluster of pneumonia associated with the 2019 novel coronavirus indicating person-to-person transmission: a study of a family cluster. *Lancet* 395, 514–523. doi: 10.1016/S0140-6736(20)30154-9
- Chang, S. L., Harding, N., Zachreson, C., Cliff, O. M., and Prokopenko, M. (2020). Modelling transmission and control of the COVID-19 pandemic in Australia. *ArXiv Prepr ArXiv200310218*. doi: 10.1038/s41467-020-19393-6
- Chapelle, O., Scholkopf, B., and Zien, A. (2009). Semi-supervised learning (chapelle, o. et al., eds.; 2006)[book reviews]. *IEEE Trans. Neural Netw.* 20:542. doi: 10.1109/TNN.2009.2015974
- Chawla, N. V., Bowyer, K. W., Hall, L. O., and Kegelmeyer, W. P. (2002). SMOTE: synthetic minority over-sampling technique. *J. Artif. Intell. Res.* 16, 321–357. doi: 10.1613/jair.953
- Chen, J., Li, K., Zhang, Z., Li, K., and Yu, P. S. (2020c). A survey on applications of artificial intelligence in fighting against covid-19. *ArXiv [Preprint]. ArXiv: 200702202*. Available online at: <https://arxiv.org/abs/2007.02202>
- Chen, J., Wu, L., Zhang, J., Zhang, L., Gong, D., Zhao, Y., et al. (2020d). Deep learning-based model for detecting 2019 novel coronavirus pneumonia on high-resolution computed tomography: a prospective study. *MedRxiv*. doi: 10.1101/2020.02.25.20021568
- Chen, J. J., Tsai, C. A., Young, J. F., and Kodell, R. L. (2005b). Classification ensembles for unbalanced class sizes in predictive toxicology. *SAR QSAR Environ. Res.* 16, 517–529. doi: 10.1080/10659360500468468
- Chen, Q., Allot, A., and Lu, Z. (2020a). Keep up with the latest coronavirus research. *Nature* 2020:193. doi: 10.1038/d41586-020-00694-1
- Chen, S., Yang, J., Yang, W., Wang, C., and Bärnighausen, T. (2020e). COVID-19 control in China during mass population movements at New Year. *Lancet* 395, 764–766. doi: 10.1016/S0140-6736(20)30421-9
- Chen, T., and Guestrin, C. (2016). “Xgboost: a scalable tree boosting system,” in *Proc. 22nd acm sigkdd Int. Conf. Knowl. Discov. data Min.* (San Francisco, CA), 785–794. doi: 10.1145/2939672.2939785
- Chen, W.-H., Strych, U., Hotez, P. J., and Bottazzi, M. E. (2020b). The SARS-CoV-2 vaccine pipeline: an overview. *Curr. Trop. Med. Rep.* 6, 1–4. doi: 10.1007/s40475-020-00201-6
- Chen, Z., Chen, J., Collins, R., Guo, Y., Peto, R., Wu, F., et al. (2011). China Kadoorie Biobank of 0.5 million people: survey methods, baseline characteristics and long-term follow-up. *Int. J. Epidemiol.* 40, 1652–1666. doi: 10.1093/ije/dyr120
- Chen, Z., Lee, L., Chen, J., Collins, R., Wu, F., Guo, Y., et al. (2005a). Cohort profile: the Kadoorie study of chronic disease in China (KSCDC). *Int. J. Epidemiol.* 34, 1243–1249. doi: 10.1093/ije/dyi174
- Cheng, C., Barceló J., Hartnett, A. S., Kubinec, R., and Messerschmidt, L. (2020). COVID-19 government response event dataset (CoronaNet v. 1.0). *Nat. Hum. Behav.* 4, 756–68. doi: 10.1038/s41562-020-0909-7
- Cho, K., Van Merriënboer, B., Bahdanau, D., and Bengio, Y. (2014). On the properties of neural machine translation: encoder-decoder approaches. *ArXiv Prepr ArXiv14091259*. doi: 10.3115/v1/W14-4012
- Chung, J., Gulcehre, C., Cho, K., and Bengio, Y. (2014). Empirical evaluation of gated recurrent neural networks on sequence modeling. *ArXiv [Preprint]. ArXiv: 14123555*. Available online at: <https://arxiv.org/abs/1412.3555>
- Chung, M., Bernheim, A., Mei, X., Zhang, N., Huang, M., Zeng, X., et al. (2020). CT imaging features of 2019 novel coronavirus (2019-nCoV). *Radiology* 2020:200230. doi: 10.1148/radiol.2020200230
- CLAIRE (2021). *The Confederation of Laboratories for Artificial Intelligence (CLAIRE) in Europe*. Available online at: <https://claire-ai.org/> (accessed March 24, 2021).
- Clark, E., Chiao, E. Y., and Amirian, E. S. (2020). Why contact tracing efforts have failed to curb coronavirus disease 2019 (covid-19) transmission in much of the united states. *Clin Infect Dis.* 2020:ciaa1155. doi: 10.1093/cid/ciaa1155
- Cohen, J. P., Morrison, P., Dao, L., Roth, K., Duong, T. Q., and Ghassemi, M. (2020). Covid-19 image data collection: prospective predictions are the future. *ArXiv [Preprint]. ArXiv: 200611988*. Available online at: <https://arxiv.org/abs/2006.11988>
- Colubri, A., Silver, T., Fradet, T., Retzepi, K., Fry, B., and Sabeti, P. (2016). Transforming clinical data into actionable prognosis models: machine-learning framework and field-deployable app to predict outcome of Ebola patients. *PLoS Negl. Trop. Dis.* 10:e0004549. doi: 10.1371/journal.pntd.0004549
- Cooper, I., Mondal, A., and Antonopoulos, C. G. (2020). A SIR model assumption for the spread of COVID-19 in different communities. *Chaos Solitons Fractals* 139:110057. doi: 10.1016/j.chaos.2020.110057
- Cortes, C., and Vapnik, V. (1995). Support-vector networks. *Mach. Learn.* 20, 273–297. doi: 10.1007/BF00994018
- Cover, T., and Hart, P. (1967). Nearest neighbor pattern classification. *IEEE Trans. Inf. Theory* 13, 21–27. doi: 10.1109/TIT.1967.1053964
- Dong, E., Du, H., and Gardner, L. (2020). An interactive web-based dashboard to track COVID-19 in real time. *Lancet Infect Dis.* 20, 533–534. doi: 10.1016/S1473-3099(20)30120-1
- Dupuis, C., De Montmollin, E., Neuville, M., Mourvillier, B., Ruckly, S., and Timsit, J. F. (2021). Limited applicability of a COVID-19 specific mortality prediction rule to the intensive care setting. *Nat. Mach. Intell.* 2020, 1–3. doi: 10.1038/s42256-020-00252-4
- Fang, Y., Zhang, H., Xie, J., Lin, M., Ying, L., Pang, P., et al. (2020). Sensitivity of chest CT for COVID-19: comparison to RT-PCR. *Radiology* 2020:200432. doi: 10.1148/radiol.2020200432
- Fong, S. J., Li, G., Dey, N., Gonzalez-Crespo, R., and Herrera-Viedma, E. (2020). Finding an accurate early forecasting model from small dataset: a case of 2019-nCoV novel coronavirus outbreak. *Int. J. Interact. Multimed. Artif. Intell.* 6, 132–140. doi: 10.9781/ijimai.2020.02.002
- Fredrikson, M., Jha, S., and Ristenpart, T. (2015). “Model inversion attacks that exploit confidence information and basic countermeasures,” in *Proc. 22nd ACM SIGSAC Conf. Comput. Commun. Secur.* (Denver, CO), 1322–1333. doi: 10.1145/281103.2813677
- Gajalakshmi, V., Peto, R., Kanimozhi, V. C., Whitlock, G., and Veeramani, D. (2007). Cohort profile: the Chennai prospective study of mortality among 500,000 adults in Tamil Nadu, South India. *Int. J. Epidemiol.* 36, 1190–1195. doi: 10.1093/ije/dym091
- Ghoshal, B., and Tucker, A. (2020). Estimating uncertainty and interpretability in deep learning for coronavirus (COVID-19) detection. *ArXiv [Preprint]. ArXiv: 200310769*. Available online at: <https://arxiv.org/abs/2003.10769>
- Giacobbe, D. R. (2020). Clinical interpretation of an interpretable prognostic model for patients with COVID-19. *Nat. Mach. Intell.* 2020:1. doi: 10.1038/s42256-020-0207-0
- Glasziou, P. P., Sanders, S., and Hoffmann, T. (2020). Waste in covid-19 research. *BMJ* 369:m1847. doi: 10.1136/bmj.m1847
- Goh, K. J., Kalimuddin, S., and Chan, K. S. (2020). Rapid progression to acute respiratory distress syndrome: review of current understanding of critical illness from coronavirus disease 2019 (COVID-19) infection. *Ann. Acad. Med. Singapore* 49, 108–118. doi: 10.47102/annals-acadmedsg.202057
- Goodfellow, I., Pouget-Abadie, J., Mirza, M., Xu, B., Warde-Farley, D., Ozair, S., et al. (2014). Generative adversarial nets. *Adv. Neural Inf. Process. Syst.* 27, 2672–2680. Available online at: <https://arxiv.org/abs/1406.2661>

- Gozes, O., Frid-Adar, M., Greenspan, H., Browning, P. D., Zhang, H., Ji, W., et al. (2020). Rapid ai development cycle for the coronavirus (covid-19) pandemic: initial results for automated detection & patient monitoring using deep learning ct image analysis. *ArXiv [Preprint]*. *ArXiv: 200305037*. Available online at: <https://arxiv.org/abs/2003.05037>
- Grant, W. B., and McDonnell, S. L. (2020). Letter in response to the article: vitamin D concentrations and COVID-19 infection in UK biobank (Hastie et al.). *Diabetes Metab. Syndr.* 15, 643–644. doi: 10.1016/j.dsx.2020.05.046
- Hack, J. J., and Papka, M. E. (2020). The US high-performance computing consortium in the fight against COVID-19. *Comput. Sci. Eng.* 22, 75–80. doi: 10.1109/MCSE.2020.3019744
- Hall, M. A., Dudek, S. M., Goodloe, R., Crawford, D. C., Pendergrass, S. A., Peissig, P., et al. (2014). Environment-wide association study (EWAS) for type 2 diabetes in the Marshfield Personalized Medicine Research Project Biobank. *Biocomput. World Sci.* 2014, 200–211. doi: 10.1142/9789814583220\_0020
- Hart, P. (1968). The condensed nearest neighbor rule (Corresp). *IEEE Trans. Inf. Theory* 14, 515–516. doi: 10.1109/TIT.1968.1054155
- Hassanien, A. E., Mahdy, L. N., Ezzat, K. A., Elmousalimi, H. H., and Ella, H. A. (2020). Automatic x-ray covid-19 lung image classification system based on multi-level thresholding and support vector machine. *MedRxiv*. doi: 10.1101/2020.03.30.20047787
- Hastie, C. E., Mackay, D. F., Ho, F., Celis-Morales, C. A., Katikireddi, S. V., Niedzwiedz, C. L., et al. (2020). Vitamin D concentrations and COVID-19 infection in UK Biobank. *Diabetes Metab. Syndr. Clin. Res. Rev.* 14, 561–565. doi: 10.1016/j.dsx.2020.04.050
- He, H., and Garcia, E. A. (2009). Learning from imbalanced data. *IEEE Trans. Knowl. Data Eng.* 21, 1263–1284. doi: 10.1109/TKDE.2008.239
- He, K., Zhang, X., Ren, S., and Sun, J. (2016). “Deep residual learning for image recognition,” in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit* (Las Vegas, NV), 770–778. doi: 10.1109/CVPR.2016.90
- Hellewell, J., Abbott, S., Gimma, A., Bosse, N. I., Jarvis, C. I., Russell, T. W., et al. (2020). Feasibility of controlling COVID-19 outbreaks by isolation of cases and contacts. *Lancet Glob. Heal.* 8, 488–496. doi: 10.1101/2020.02.08.20021162
- Hemdan, E. E.-D., Shouman, M. A., and Karar, M. E. (2020). Covidx-net: a framework of deep learning classifiers to diagnose covid-19 in x-ray images. *ArXiv [Preprint]*. *ArXiv: 200311055*. Available online at: <https://arxiv.org/abs/2003.11055>
- Hitaj, B., Ateniense, G., and Perez-Cruz, F. (2017). “Deep models under the GAN: information leakage from collaborative deep learning,” in *Proc. 2017 ACM SIGSAC Conf. Comput. Commun. Secur.* (Dallas, TX), 603–618. doi: 10.1145/3133956.3134012
- Hochreiter, S., and Schmidhuber, J. (1997). Long short-term memory. *Neural Comput.* 9, 1735–1780. doi: 10.1162/neco.1997.9.8.1735
- Hope, T., Portenoy, J., Vasani, K., Borchardt, J., Horvitz, E., Weld, D. S., et al. (2020). SciSight: combining faceted navigation and research group detection for COVID-19 exploratory scientific search. *ArXiv Prepr ArXiv200512668*. doi: 10.1101/2020.05.23.112284
- Hopfield, J. J. (1988). Artificial neural networks. *IEEE Circuits Devices Mag.* 4, 3–10. doi: 10.1109/101.8118
- Hosmer, D. W. Jr., Lemeshow, S., and Sturdivant, R. X. (2013). *Applied Logistic Regression*. Vol. 398. Hoboken, NJ: John Wiley & Sons. doi: 10.1002/9781118548387
- Hu, Z., Ge, Q., Li, S., Boerwinkle, E., Jin, L., and Xiong, M. (2020). Forecasting and evaluating multiple interventions for COVID-19 worldwide. *Front. Artif. Intell.* 3:41. doi: 10.3389/frai.2020.00041
- Huang, C., Wang, Y., Li, X., Ren, L., Zhao, J., Hu, Y., et al. (2020a). Clinical features of patients infected with 2019 novel coronavirus in Wuhan, China. *Lancet* 395, 497–506. doi: 10.1016/S0140-6736(20)30183-5
- Huang, C.-J., Chen, Y.-H., Ma, Y., and Kuo, P.-H. (2020b). Multiple-input deep convolutional neural network model for covid-19 forecasting in china. *MedRxiv*. doi: 10.1101/2020.03.23.20041608
- Imran, A., Posokhova, I., Qureshi, H. N., Masood, U., Riaz, S., Ali, K., et al. (2020). AI4COVID-19: AI enabled preliminary diagnosis for COVID-19 from cough samples via an app. *ArXiv Prepr ArXiv200401275*. doi: 10.1016/j.imu.2020.100378
- Iqbal, M. Z., and Faiz, M. F. I. (2020). Active Surveillance for COVID-19 through artificial intelligence using concept of real-time speech-recognition mobile application to analyse cough sound. *arXiv*. doi: 10.31219/osf.io/cev6x
- Isham, V., and Medley, G. (1996). *Models for Infectious Human Diseases: Their Structure and Relation to Data*. Vol. 6. Cambridge: Cambridge University Press. doi: 10.1017/CBO9780511662935
- Ishmaev, G., Dennis, M., and van den Hoven, M. J. (2021). Ethics in the COVID-19 pandemic: myths, false dilemmas, and moral overload. *Ethics Inf. Technol.* 2021, 1–16. doi: 10.1007/s10676-020-09568-6
- Jain, A. K., Mao, J., and Mohiuddin, K. M. (1996). Artificial neural networks: a tutorial. *Computer* 29, 31–44. doi: 10.1109/2.485891
- Jamshidi, M., Lalbakhsh, A., Talla, J., Peroutka, Z., Hadjilooei, F., Lalbakhsh, P., et al. (2020). Artificial intelligence and COVID-19: deep learning approaches for diagnosis and treatment. *IEEE Access* 8, 109581–109595. doi: 10.1109/ACCESS.2020.3001973
- Jana, S., and Bhaumik, P. (2020). A multivariate spatiotemporal spread model of COVID-19 using ensemble of ConvLSTM networks. *MedRxiv* (2020).
- Jang, J.-S. (1993). ANFIS: adaptive-network-based fuzzy inference system. *IEEE Trans. Syst. Man. Cybern.* 23, 665–685. doi: 10.1109/21.256541
- Janiaud, P., Axfors, C., Van't Hooft, J., Saccolotto, R., Agarwal, A., Appenzeller-Herzog, C., et al. (2020). The worldwide clinical trial research response to the COVID-19 pandemic-the first 100 days. *F1000Research* 9:2. doi: 10.12688/f1000research.26707.2
- Jewell, N. P., Lewnard, J. A., and Jewell, B. L. (2020). Predictive mathematical models of the COVID-19 pandemic: underlying principles and value of projections. *JAMA* 323, 1893–1894. doi: 10.1001/jama.2020.6585
- Jiang, X., Coffee, M., Bari, A., Wang, J., Jiang, X., Huang, J., et al. (2020a). Towards an artificial intelligence framework for data-driven prediction of coronavirus clinical severity. *C Comput. Mater. Contin.* 63, 537–551. doi: 10.32604/cmc.2020.010691
- Jiang, Z., Hu, M., Gao, Z., Fan, L., Dai, R., Pan, Y., et al. (2020b). Detection of respiratory infections using RGB-infrared sensors on portable device. *IEEE Sens. J.* 20, 13674–13681. doi: 10.1109/JSEN.2020.3004568
- Jimenez-Solem, E., Petersen, T. S., Hansen, C., Hansen, C., Lioma, C., Igel, C., et al. (2020). Developing and validating COVID-19 adverse outcome risk prediction models from a bi-national European Cohort of 5,594 patients. *MedRxiv*. doi: 10.1101/2020.10.06.20207209
- Jin, C., Chen, W., Cao, Y., Xu, Z., Zhang, X., Deng, L., et al. (2020b). Development and evaluation of an AI system for COVID-19 diagnosis. *Nat. Commun.* 11:20039834. doi: 10.1101/2020.03.20.20039834
- Jin, Y., Yang, H., Ji, W., Wu, W., Chen, S., Zhang, W., et al. (2020a). Virology, epidemiology, pathogenesis, and control of COVID-19. *Viruses* 12:372. doi: 10.3390/v12040372
- John, M., and Shaiba, H. (2019). Main factors influencing recovery in MERS Co-V patients using machine learning. *J. Infect. Public Health* 12, 700–704. doi: 10.1016/j.jiph.2019.03.020
- Johnson, K. W., Shameer, K., Glicksberg, B. S., Readhead, B., Sengupta, P. P., Björkregren, J. L. M., et al. (2017). Enabling precision cardiology through multiscale biology and systems medicine. *JACC Basic Transl. Sci.* 2, 311–327. doi: 10.1016/j.jacbts.2016.11.010
- Kaiser Permanente (2020). *Kaiser Permanente: Research Program on Genes, Environment and Health*. Available online at: <https://divisionofresearch.kaiserpermanente.org/genetics/rpgeh/rpgeabout> (accessed November 1, 2020).
- Kaissis, G. A., Makowski, M. R., Rückert, D., and Braren, R. F. (2020). Secure, privacy-preserving and federated machine learning in medical imaging. *Nat. Mach. Intell.* 2, 305–311. doi: 10.1038/s42256-020-0186-1
- Kandel, N., Chungong, S., Omaar, A., Xing, J. (2020). Health security capacities in the context of COVID-19 outbreak: an analysis of International Health Regulations annual report data from 182 countries. *Lancet* 395, 1047–1053. doi: 10.1016/S0140-6736(20)30553-5
- Kang, H., Xia, L., Yan, F., Wan, Z., Shi, F., Yuan, H., et al. (2020). Diagnosis of coronavirus disease 2019 (covid-19) with structured latent multi-view representation learning. *IEEE Trans. Med. Imaging.* 39, 2606–2614. doi: 10.1109/TMI.2020.2992546
- Kanne, J. P. (2020). Chest CT findings in 2019 novel coronavirus (2019-nCoV) infections from Wuhan, China: key points for the radiologist. *Radiology* 2020:200241. doi: 10.1148/radiol.202002041
- Kanungo, T., Mount, D. M., Netanyahu, N. S., Piatko, C. D., Silverman, R., and Wu, A. Y. (2002). An efficient k-means clustering algorithm: analysis

- and implementation. *IEEE Trans. Pattern Anal. Mach. Intell.* 24, 881–892. doi: 10.1109/TPAMI.2002.1017616
- Karako, K., Song, P., Chen, Y., and Tang, W. (2020). Analysis of COVID-19 infection spread in Japan based on stochastic transition model. *Biosci. Trends.* 14, 134–138. doi: 10.5582/bst.2020.01482
- Kavadi, D. P., Patan, R., Ramachandran, M., and Gandomi, A. H. (2020). Partial derivative non-linear global pandemic machine learning prediction of covid 19. *Chaos Solitons Fractals* 139:110056. doi: 10.1016/j.chaos.2020.110056
- Kenneth, C. Y., and So, H.-C. (2020). Uncovering clinical risk factors and prediction of severe COVID-19: a machine learning approach based on UK Biobank data. *MedRxiv*. doi: 10.1101/2020.09.18.20197319
- Kissler, S. M., Tedijanto, C., Lipsitch, M., and Grad, Y. (2020). Social distancing strategies for curbing the COVID-19 epidemic. *MedRxiv*. doi: 10.1101/2020.03.22.20041079
- Kong, W.-H., Li, Y., Peng, M.-W., Kong, D.-G., Yang, X.-B., Wang, L., et al. (2020). SARS-CoV-2 detection in patients with influenza-like illness. *Nat. Microbiol.* 5, 675–678. doi: 10.1038/s41564-020-0713-1
- Koo, J. R., Cook, A. R., Park, M., Sun, Y., Sun, H., Lim, J. T., et al. (2020). Interventions to mitigate early spread of SARS-CoV-2 in Singapore: a modelling study. *Lancet Infect. Dis.* 20, 678–688. doi: 10.1016/S1473-3099(20)30162-6
- Kooraki, S., Hosseiny, M., Myers, L., and Gholamrezaezhad, A. (2020). Coronavirus (COVID-19) outbreak: what the department of radiology should know. *J. Am. Coll. Radiol.* 17, 447–451. doi: 10.1016/j.jacr.2020.02.008
- Kotsiantis, S. B., Zaharakis, I., and Pintelas, P. (2007). Supervised machine learning: a review of classification techniques. *Emerg. Artif. Intell. Appl. Comput. Eng.* 160, 3–24. doi: 10.1007/s10462-007-9052-3
- Kretzschmar, M., Rozhnova, G., and van Boven, M. (2020b). Isolation and contact tracing can tip the scale to containment of COVID-19 in populations with social distancing. *SSRN 3562458*. doi: 10.2139/ssrn.3562458
- Kretzschmar, M. E., Rozhnova, G., Bootsma, M. C. J., van Boven, M., van de Wijgert, J. H. H. M., and Bonten, M. J. M. (2020a). Impact of delays on effectiveness of contact tracing strategies for COVID-19: a modelling study. *Lancet Public Heal.* 5, e452–e459. doi: 10.1016/S2468-2667(20)30157-2
- Kricka, L. J., Polevikov, S., Park, J. Y., Fortina, P., Bernardini, S., Satchkov, D., et al. (2020). Artificial intelligence-powered search tools and resources in the fight against COVID-19. *Ejifcc* 31:106.
- Krizhevsky, A., Sutskever, I., and Hinton, G. E. (2017). ImageNet classification with deep convolutional neural networks. *Commun. ACM* 60, 84–90. doi: 10.1145/3065386
- Kubat, M., and Matwin, S. (1997). Addressing the curse of imbalanced training sets: one-sided selection. *ICML 97*, 179–186.
- Kumar, A., Gupta, P. K., and Srivastava, A. (2020). A review of modern technologies for tackling COVID-19 pandemic. *Diabetes Metab Syndr Clin Res Rev.* 14, 569–573. doi: 10.1016/j.dsx.2020.05.008
- Lallie, H. S., Shepherd, L. A., Nurse, J. R. C., Erola, A., Epiphaniou, G., Maple, C., et al. (2021). Cyber security in the age of covid-19: a timeline and analysis of cyber-crime and cyber-attacks during the pandemic. *Comput. Secur.* 2021:102248. doi: 10.1016/j.cose.2021.102248
- Lalmuanawma, S., Hussain, J., and Chhakhuak, L. (2020). Applications of machine learning and artificial intelligence for Covid-19 (SARS-CoV-2) pandemic: a review. *Chaos Solitons Fractals* 2020:110059. doi: 10.1016/j.chaos.2020.110059
- Laurikkala, J. (2001). “Improving identification of difficult small classes by balancing class distribution,” in *Conf. Artif. Intell. Med. Eur* (Berlin), 63–66. doi: 10.1007/3-540-48229-6\_9
- LeCun, Y., Bengio, Y., and Hinton, G. (2015). Deep learning. *Nature* 521, 436–444. doi: 10.1038/nature14539
- Lee, J., Sun, J., Wang, F., Wang, S., Jun, C.-H., and Jiang, X. (2018). Privacy-preserving patient similarity learning in a federated environment: development and analysis. *JMIR Med. Informatics* 6:e20. doi: 10.2196/medinform.7744
- Lewis, D. (2020). Why many countries failed at COVID contact-tracing-but some got it right. *Nature* 588, 384–387. doi: 10.1038/d41586-020-03518-4
- Li, L., Qin, L., Xu, Z., Yin, Y., Wang, X., Kong, B., et al. (2020b). Artificial intelligence distinguishes COVID-19 from community acquired pneumonia on chest CT. *Radiology*. 2020:200905. doi: 10.1148/radiol.2020.200905
- Li, L., Zhang, Q., Wang, X., Zhang, J., Wang, T., Gao, T.-L., et al. (2020a). Characterizing the propagation of situational information in social media during covid-19 epidemic: a case study on weibo. *IEEE Trans. Comput. Soc. Syst.* 7, 556–562. doi: 10.1109/TCSS.2020.2980007
- Li, M., Lei, P., Zeng, B., Li, Z., Yu, P., Fan, B., et al. (2020c). Coronavirus disease (COVID-19): spectrum of CT findings and temporal progression of the disease. *Acad. Radiol.* 27, 603–608. doi: 10.1016/j.acra.2020.03.003
- Li, T., Sahu, A. K., Talwalkar, A., and Smith, V. (2020d). Federated learning: challenges, methods, and future directions. *IEEE Signal Process Mag.* 37, 50–60. doi: 10.1109/MSP.2020.2975749
- Liaw, A., and Wiener, M. (2002). Classification and regression by randomForest. *R News* 2, 18–22. Available online at: <https://cogsci.northwestern.edu/cbmg/LiawAndWiener2002.pdf>
- Lu, H., and Wang, M. (2019). RL4health: crowdsourcing reinforcement learning for knee replacement pathway optimization. *ArXiv [Preprint]*. *ArXiv: 190601407*. Available online at: <https://arxiv.org/abs/1906.01407>
- Lu, R., Wu, X., Wan, Z., Li, Y., Zuo, L., Qin, J., et al. (2020a). Development of a novel reverse transcription loop-mediated isothermal amplification method for rapid detection of SARS-CoV-2. *Viol. Sin.* 2020:1. doi: 10.3390/ijms21082826
- Lu, R., Zhao, X., Li, J., Niu, P., Yang, B., Wu, H., et al. (2020b). Genomic characterisation and epidemiology of 2019 novel coronavirus: implications for virus origins and receptor binding. *Lancet* 395, 565–574. doi: 10.1016/S0140-6736(20)30251-8
- Maghdid, H. S., Asaad, A. T., Ghafoor, K. Z., Sadiq, A. S., and Khan, M. K. (2020a). Diagnosing COVID-19 pneumonia from X-ray and CT images using deep learning and transfer learning algorithms. *ArXiv Prepr ArXiv200400038*. doi: 10.1117/12.2588672
- Maghdid, H. S., Ghafoor, K. Z., Sadiq, A. S., Curran, K., and Rabie, K. (2020b). A novel ai-enabled framework to diagnose coronavirus covid 19 using smartphone embedded sensors: design study. *ArXiv Prepr ArXiv200307434*. doi: 10.1109/IRI49571.2020.00033
- Mahler, T., Nissim, N., Shalom, E., Goldenberg, I., Hassman, G., Makori, A., et al. (2018). Know your enemy: characteristics of cyber-attacks on medical imaging devices. *ArXiv [Preprint]*. *ArXiv: 180105583*. Available online at: <https://arxiv.org/abs/1801.05583>
- Martin, A., Nateqi, J., Gruarin, S., Munsch, N., Abdarrahmane, I., Zobel, M., et al. (2020). An artificial intelligence-based first-line defence against COVID-19: digitally screening citizens for risks via a chatbot. *Sci. Rep.* 10, 1–7. doi: 10.1038/s41598-020-75912-x
- Mauger, C., Gilbert, K., Lee, A. M., Sanghvi, M. M., Aung, N., Fung, K., et al. (2019). Right ventricular shape and function: cardiovascular magnetic resonance reference morphology and biventricular risk factor morphometrics in UK Biobank. *J. Cardiovasc. Magn. Reson.* 21:41. doi: 10.1186/s12968-019-0551-6
- May, R. M., and Anderson, R. M. (1979). Population biology of infectious diseases: part I. *Nature* 280, 455–461. doi: 10.1038/280455a0
- Mena-Lorcat, J., and Hethcote, H. W. (1992). Dynamic models of infectious diseases as regulators of population sizes. *J. Math. Biol.* 30, 693–716. doi: 10.1007/BF00173264
- Moriyama, M., Hugentobler, W. J., and Iwasaki, A. (2020). Seasonality of respiratory viral infections. *Annu. Rev. Virol.* 7:22445. doi: 10.1146/annurev-virology-012420-022445
- Muthuppalaniappan, M., and Stevenson, K. (2021). Healthcare cyber-attacks and the COVID-19 pandemic: an urgent threat to global health. *Int. J. Qual. Heal. Care* 33:mzaa117. doi: 10.1093/intqhc/mzaa117
- Narin, A., Kaya, C., and Pamuk, Z. (2020). Automatic detection of coronavirus disease (covid-19) using x-ray images and deep convolutional neural networks. *ArXiv [Preprint]*. *ArXiv: 200310849*. Available online at: <https://arxiv.org/abs/2003.10849>
- Naudé, W. (2020). Artificial intelligence vs. COVID-19: limitations, constraints and pitfalls. *Ai Soc.* 2020:1. doi: 10.1007/s00146-020-00978-0
- NCATS-US (2021). *National COVID Cohort Collaborative (N3C)*. Available online at: <https://ncats.nih.gov/n3c> (accessed March 24, 2021).
- Ndairou, F., Area, I., Nieto, J. J., and Torres, D. F. M. (2020). Mathematical modeling of COVID-19 transmission dynamics with a case study of Wuhan. *Chaos Solitons Fractals* 2020:109846. doi: 10.1016/j.chaos.2020.109846
- Nguyen, T. T. (2020). Artificial intelligence in the battle against coronavirus (COVID-19): a survey and future research directions. *ArXiv Prepr ArXiv*. doi: 10.36227/techrxiv.12743933

- NHSX. (2021). *National COVID-19 Chest Image Database (NCCID)*. Available online at: <https://nhsx.github.io/covid-chest-imaging-database/index.html> (accessed March 24, 2021).
- Oh, S.-H. (2011). Error back-propagation algorithm for classification of imbalanced data. *Neurocomputing* 74, 1058–1061. doi: 10.1016/j.neucom.2010.11.024
- Ozturk, T., Talo, M., Yildirim, E. A., Baloglu, U. B., Yildirim, O., and Acharya, U. R. (2020). Automated detection of COVID-19 cases using deep neural networks with X-ray images. *Comput. Biol. Med.* 2020:103792. doi: 10.1016/j.compbiomed.2020.103792
- Pan, F., Ye, T., Sun, P., Gui, S., Liang, B., Li, L., et al. (2020). Time course of lung changes on chest CT during recovery from 2019 novel coronavirus (COVID-19) pneumonia. *Radiology* 2020:200370. doi: 10.1148/radiol.2020200370
- Park, Y. J., Choe, Y. J., Park, O., Park, S. Y., Kim, Y.-M., Kim, J., et al. (2020). Contact tracing during coronavirus disease outbreak, South Korea, 2020. *Emerg. Infect. Dis.* 26, 2465–2468. doi: 10.3201/eid2610.201315
- Pascual, S., Bonafonte, A., and Serra, J. (2017). SEGAN: speech enhancement generative adversarial network. *ArXiv Prepr ArXiv170309452*. doi: 10.21437/Interspeech.2017-1428
- Pastor-Satorras, R., Castellano, C., Van Mieghem, P., and Vespignani, A. (2015). Epidemic processes in complex networks. *Rev. Mod. Phys.* 87:925. doi: 10.1103/RevModPhys.87.925
- Patel, S. S., Webster, R. K., Greenberg, N., Weston, D., and Brooks, S. K. (2020). Research fatigue in COVID-19 pandemic and post-disaster research: causes, consequences and recommendations. *Disaster Prev. Manag. An Int. J.* 29, 445–455. doi: 10.1108/DPM-05-2020-0164
- Pereira, N. L., Ahmad, F., Cummins, N. W., Byku, M., Morris, A. A., Owens, A., et al. (2020). COVID-19: understanding inter-individual variability and implications for precision medicine. *Mayo Clin. Proc.* 96, 446–463. doi: 10.1016/j.mayocp.2020.11.024
- Petersen, S. E., Abdulkareem, M., and Leiner, T. (2019). Artificial intelligence will transform cardiac imaging—opportunities and challenges. *Front. Cardiovasc. Med.* 6:133. doi: 10.3389/fcvm.2019.00133
- Petersen, S. E., Aung, N., Sanghvi, M. M., Zemrak, F., Fung, K., Paiva, J. M., et al. (2017). Reference ranges for cardiac structure and function using cardiovascular magnetic resonance (CMR) in Caucasians from the UK Biobank population cohort. *J. Cardiovasc. Magn. Reson.* 19:18. doi: 10.1186/s12968-017-0327-9
- Phua, J., Weng, L., Ling, L., Egi, M., Lim, C.-M., Divatia, J. V., et al. (2020). Intensive care management of coronavirus disease 2019 (COVID-19): challenges and recommendations. *Lancet Respir. Med.* 8, 506–517. doi: 10.1016/S2213-2600(20)30161-2
- Pourhomayoun, M., and Shakibi, M. (2020). Predicting mortality risk in patients with COVID-19 using artificial intelligence to help medical decision-making. *MedRxiv*. doi: 10.1101/2020.03.30.20047308
- Prudêncio, M., and Costa, J. C. (2020). Research funding after COVID-19. *Nat. Microbiol.* 5:986. doi: 10.1038/s41564-020-0768-z
- Qi, X., Jiang, Z., Yu, Q., Shao, C., Zhang, H., Yue, H., et al. (2020). Machine learning-based CT radiomics model for predicting hospital stay in patients with pneumonia associated with SARS-CoV-2 infection: a multicenter study. *MedRxiv*. doi: 10.1101/2020.02.29.20029603
- Qian, X., Ren, R., Wang, Y., Guo, Y., Fang, J., Wu, Z.-D., et al. (2020). Fighting against the common enemy of COVID-19: a practice of building a community with a shared future for mankind. *Infect Dis. Poverty* 9, 1–6. doi: 10.1186/s40249-020-00650-1
- Quanjel, M. J. R., van Holten, T. C., der Vliet, P. C., Wielaard, J., Karakaya, B., Söhne, M., et al. (2020). Replication of a mortality prediction model in Dutch patients with COVID-19. *Nat. Mach. Intell.* 2020, 1–2. doi: 10.1038/s42256-020-00253-3
- Raghu, M., and Schmidt, E. (2020). A survey of deep learning for scientific discovery. *ArXiv [Preprint]*. *ArXiv: 200311755*. Available online at: <https://arxiv.org/abs/2003.11755>
- Rasheed, J., Jamil, A., Hameed, A. A., Aftab, U., Aftab, J., Shah, S. A., et al. (2020). A survey on artificial intelligence approaches in supporting frontline workers and decision makers for COVID-19 pandemic. *Chaos Solitons Fractals* 2020:110337. doi: 10.1016/j.chaos.2020.110337
- Ren, S., He, K., Girshick, R., and Sun, J. (2016). Faster r-cnn: towards real-time object detection with region proposal networks. *IEEE Trans. Pattern Anal. Mach. Intell.* 39, 1137–1149. doi: 10.1109/TPAMI.2016.2577031
- Riad, M. H., Sekamatte, M., Ocom, F., Makumbi, I., and Scoglio, C. M. (2019). Risk assessment of Ebola virus disease spreading in Uganda using a two-layer temporal network. *Sci. Rep.* 9, 1–17. doi: 10.1038/s41598-019-52501-1
- Ribeiro, M. H. D. M., da Silva, R. G., Mariani, V. C., and dos Santos Coelho, L. (2020). Short-term forecasting COVID-19 cumulative confirmed cases: perspectives for Brazil. *Chaos Solitons Fractals* 2020:109853. doi: 10.1016/j.chaos.2020.109853
- Rieke, N., Hancox, J., Li, W., Milletari, F., Roth, H., Albarqouni, S., et al. (2020). The future of digital health with federated learning. *ArXiv Prepr ArXiv200308119*. doi: 10.1038/s41746-020-00323-1
- Robinson, R., Valindria, V. V., Bai, W., Oktay, O., Kainz, B., Suzuki, H., et al. (2019). Automated quality control in image segmentation: application to the UK Biobank cardiovascular magnetic resonance imaging study. *J. Cardiovasc. Magn. Reson.* 21, 1–14. doi: 10.1186/s12968-019-0523-x
- Roden, D. M., Pulley, J. M., Basford, M. A., Bernard, G. R., Clayton, E. W., Balsler, J. R., et al. (2008). Development of a large-scale de-identified DNA biobank to enable personalized medicine. *Clin. Pharmacol. Ther.* 84, 362–369. doi: 10.1038/clpt.2008.89
- Rodriguez, F., Scheinker, D., and Harrington, R. A. (2018). Promise and perils of big data and artificial intelligence in clinical medicine and biomedical research. *Circ. Res.* 123, 1282–1284. doi: 10.1161/CIRCRESAHA.118.314119
- Ronneberger, O., Fischer, P., and Brox, T. (2015). U-net: convolutional networks for biomedical image segmentation. *Lect. Notes Comput. Sci.* 9351, 234–241. Available online at: <https://arxiv.org/abs/1505.04597>
- Rothan, H. A., and Byrareddy, S. N. (2020). The epidemiology and pathogenesis of coronavirus disease (COVID-19) outbreak. *J. Autoimmun.* 2020:102433. doi: 10.1016/j.jaut.2020.102433
- Ruiz Estrada, M. A. (2020). The uses of drones in case of massive epidemics contagious diseases relief humanitarian aid: Wuhan-COVID-19 crisis. *SSRN 3546547*. doi: 10.2139/ssrn.3546547
- Russakovsky, O., Deng, J., Su, H., Krause, J., Satheesh, S., Ma, S., et al. (2015). Imagenet large scale visual recognition challenge. *Int. J. Comput. Vis.* 115, 211–252. doi: 10.1007/s11263-015-0816-y
- Rutten-Jacobs, L. C. A., Larsson, S. C., Malik, R., Rannikmäe, K., Sudlow, C. L., Dichgans, M., et al. (2018). Genetic risk, incident stroke, and the benefits of adhering to a healthy lifestyle: cohort study of 306 473 UK Biobank participants. *BMJ* 363:k4168. doi: 10.1136/bmj.k4168
- Said, M. A., Verweij, N., and van der Harst, P. (2018). Associations of combined genetic and lifestyle risks with incident cardiovascular disease and diabetes in the UK Biobank Study. *JAMA Cardiol.* 3, 693–702. doi: 10.1001/jamacardio.2018.1717
- Salathé M., Althaus, C. L., Neher, R., Stringhini, S., Hodcroft, E., Fellay, J., et al. (2020). COVID-19 epidemic in Switzerland: on the importance of testing, contact tracing and isolation. *Swiss Med. Wkly* 150:w20225. doi: 10.4414/smww.2020.20225
- Sameni, R. (2020). Mathematical modeling of epidemic diseases; a case study of the COVID-19 coronavirus. *ArXiv [Preprint]*. *ArXiv: 200311371*. Available online at: <https://arxiv.org/abs/2003.11371>
- Sattar, N., Ho, F. K., Gill, J. M. R., Ghouri, N., Gray, S. R., Celis-Morales, C. A., et al. (2020). BMI and future risk for COVID-19 infection and death across sex, age and ethnicity: preliminary findings from UK biobank. *Diabetes Metab. Syndr. Clin. Res. Rev.* 14, 1149–1151. doi: 10.1016/j.dsx.2020.06.060
- Schuller, B. W., Schuller, D. M., Qian, K., Liu, J., Zheng, H., and Li, X. (2020). Covid-19 and computer audition: an overview on what speech and sound analysis could contribute in the SARS-CoV-2 Corona crisis. *ArXiv Prepr ArXiv200311117*. doi: 10.3389/fdgth.2021.564906
- Schwartz, D. A. (2020). An analysis of 38 pregnant women with COVID-19, their newborn infants, and maternal-fetal transmission of SARS-CoV-2: maternal coronavirus infections and pregnancy outcomes. *Arch. Pathol. Lab. Med.* 144, 799–805. doi: 10.5858/arpa.2020-0901-SA
- Shen, D., Wu, G., and Suk, H.-I. (2017). Deep learning in medical image analysis. *Annu. Rev. Biomed. Eng.* 19, 221–248. doi: 10.1146/annurev-bioeng-071516-044442
- Shi, F., Wang, J., Shi, J., Wu, Z., Wang, Q., Tang, Z., et al. (2020a). Review of artificial intelligence techniques in imaging data acquisition,

- segmentation and diagnosis for covid-19. *IEEE Rev. Biomed. Eng.* 14:2987975. doi: 10.1109/RBME.2020.2987975
- Shi, F., Xia, L., Shan, F., Wu, D., Wei, Y., Yuan, H., et al. (2020b). Large-scale screening of covid-19 from community acquired pneumonia using infection size-aware classification. *ArXiv Prepr ArXiv200309860*. doi: 10.1088/1361-6560/abe838
- Shi, W., Peng, X., Liu, T., Cheng, Z., Lu, H., Yang, S., et al. (2020c). Deep learning-based quantitative computed tomography model in predicting the severity of COVID-19: a retrospective study in 196 patients. *Ann. Transl. Med.* 9:216. doi: 10.2139/ssrn.3546089
- Shivade, C., Raghavan, P., Fosler-Lussier, E., Embi, P. J., Elhadad, N., Johnson, S. B., et al. (2014). A review of approaches to identifying patient phenotype cohorts using electronic health records. *J. Am. Med. Informat. Assoc.* 21, 221–230. doi: 10.1136/amiajnl-2013-001935
- Shortliffe, E. H., and Sepúlveda, M. J. (2018). Clinical decision support in the era of artificial intelligence. *JAMA* 320, 2199–2200. doi: 10.1001/jama.2018.17163
- Siettos, C. I., and Russo, L. (2013). Mathematical modeling of infectious disease dynamics. *Virulence* 4, 295–306. doi: 10.4161/viru.24041
- SoBigData.eu (2021). *The Project SoBigData*. Available online at: <http://project.sobigdata.eu/> (accessed March 24, 2021).
- Song, Y., Zheng, S., Li, L., Zhang, X., Zhang, X., Huang, Z., et al. (2020). Deep learning enables accurate diagnosis of novel coronavirus (COVID-19) with CT images. *MedRxiv*. doi: 10.1109/TCBB.2021.3065361
- Suinesiaputra, A., Sanghvi, M. M., Aung, N., Paiva, J. M., Zemrak, F., Fung, K., et al. (2018). Fully-automated left ventricular mass and volume MRI analysis in the UK Biobank population cohort: evaluation of initial results. *Int. J. Cardiovasc. Imaging* 34, 281–291. doi: 10.1007/s10554-017-1225-9
- Suri, J. S., Puvvula, A., Biswas, M., Majhail, M., Saba, L., Faa, G., et al. (2020). COVID-19 pathways for brain and heart injury in comorbidity patients: a role of medical imaging and artificial intelligence-based COVID severity classification: a review. *Comput. Biol. Med.* 2020:103960. doi: 10.1016/j.compbiomed.2020.103960
- Sutton, R. S., and Barto, A. G. (1998). *Introduction to Reinforcement Learning*. Vol. 135. Cambridge: MIT press.
- Szegedy, C., Vanhoucke, V., Ioffe, S., Shlens, J., and Wojna, Z. (2016). “Rethinking the inception architecture for computer vision,” in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, 2818–2826. doi: 10.1109/CVPR.2016.308
- Tang, Z., Zhao, W., Xie, X., Zhong, Z., Shi, F., Liu, J., et al. (2020). Severity assessment of coronavirus disease 2019 (COVID-19) using quantitative features from chest CT images. *ArXiv Prepr ArXiv200311988*. doi: 10.1088/1361-6560/abbf9e
- Team, E. E. (2020). Note from the editors: World Health Organization declares novel coronavirus (2019-nCoV) sixth public health emergency of international concern. *Eurosurveillance* 25:200131e. doi: 10.2807/1560-7917.ES.2020.25.5.200131e
- Thevarajan, I., Nguyen, T. H. O., Koutsakos, M., Druce, J., Caly, L., van de Sandt, C. E., et al. (2020). Breadth of concomitant immune responses prior to patient recovery: a case report of non-severe COVID-19. *Nat. Med.* 26, 453–455. doi: 10.1038/s41591-020-0819-2
- Tibshirani, R. (1996). Regression shrinkage and selection via the lasso. *J. R. Stat. Soc. Ser. B* 58, 267–288. doi: 10.1111/j.2517-6161.1996.tb02080.x
- Toh, C., and Brody, J. P. (2020). Evaluation of a genetic risk score for severity of COVID-19 using human chromosomal-scale length variation. *Hum. Genom.* 14, 1–5. doi: 10.1186/s40246-020-00288-y
- Udugama, B., Kadhiresan, P., Kozłowski, H. N., Malekjahani, A., Osborne, M., Li, V. Y. C., et al. (2020). Diagnosing COVID-19: the disease and tools for detection. *ACS Nano* 14, 3822–3835. doi: 10.1021/acsnano.0c02624
- UKCDR (2020). *COVID-19 Research Project Tracker by UKCDR & GloPID-R*. *UK Collab Dev Res*. Available online at: <https://www.ukcdr.org.uk/covid-circle/covid-19-research-project-tracker/> (accessed January 1, 2021).
- Vaishya, R., Haleem, A., Vaish, A., and Javaid, M. (2020a). Emerging technologies to combat the COVID-19 pandemic. *J. Clin. Exp. Hepatol.* 10, 409–411. doi: 10.1016/j.jceh.2020.04.019
- Vaishya, R., Javaid, M., Khan, I. H., and Haleem, A. (2020b). Artificial Intelligence (AI) applications for COVID-19 pandemic. *Diabetes Metab. Syndr. Clin. Res. Rev.* 14, 337–339. doi: 10.1016/j.dsx.2020.04.012
- Vynnycky, E., and White, R. (2010). *An Introduction to Infectious Disease Modelling*. Oxford: OUP.
- Walls, A. C., Park, Y.-J., Tortorici, M. A., Wall, A., McGuire, A. T., and Veesler, D. (2020). Structure function and antigenicity of the SARS-CoV-2 spike glycoprotein. *Cell* 181, 281–292. doi: 10.1016/j.cell.2020.02.058
- Wan, Y., Shang, J., Graham, R., Baric, R. S., and Li, F. (2020). Receptor recognition by novel coronavirus from Wuhan: an analysis based on decade-long structural studies of SARS. *J. Virol.* 2020, JVI00127–00120. doi: 10.1128/JVI.00127-20
- Wang, C. J., Ng, C. Y., and Brook, R. H. (2020b). Response to COVID-19 in Taiwan: big data analytics, new technology, and proactive testing. *JAMA*. 2020:10.1001/jama.2020.3151. doi: 10.1001/jama.2020.3151
- Wang, J., Zhang, F., Zeng, J., Wu, Y., Kemper, K. E., Xue, A., et al. (2019). Genotype-by-environment interactions inferred from genetic effects on phenotypic variability in the UK Biobank. *Sci. Adv.* 5:eaw3538. doi: 10.1126/sciadv.aaw3538
- Wang, L., Lin, Z. Q., and Wong, A. (2020d). Covid-net: a tailored deep convolutional neural network design for detection of covid-19 cases from chest x-ray images. *Sci. Rep.* 10, 1–12. doi: 10.1038/s41598-020-76550-z
- Wang, S., Kang, B., Ma, J., Zeng, X., Xiao, M., Guo, J., et al. (2020c). A deep learning algorithm using CT images to screen for Corona Virus Disease (COVID-19). *MedRxiv*. doi: 10.1101/2020.02.14.20023028
- Wang, X., Deng, X., Fu, Q., Zhou, Q., Feng, J., Ma, H., et al. (2020e). A weakly-supervised framework for COVID-19 classification and lesion localization from chest CT. *IEEE Trans. Med. Imaging*. 39:2995963. doi: 10.1109/TMI.2020.2995965
- Wang, X., Peng, Y., Lu, L., Lu, Z., Bagheri, M., Summers, R. M., et al. (2017). “Chestx-ray8: Hospital-scale chest x-ray database and benchmarks on weakly-supervised classification and localization of common thorax diseases,” in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, 2097–2106. doi: 10.1109/CVPR.2017.369
- Wang, Y., Hu, M., Li, Q., Zhang, X.-P., Zhai, G., and Yao, N. (2020f). Abnormal respiratory patterns classifier may contribute to large-scale screening of people infected with COVID-19 in an accurate and unobtrusive manner. *ArXiv [Preprint]*. *ArXiv: 200205534*. Available online at: <https://arxiv.org/abs/2002.05534>
- Wang, Y., Wang, Y., Chen, Y., and Qin, Q. (2020a). Unique epidemiological and clinical features of the emerging 2019 novel coronavirus pneumonia (COVID-19) implicate special control measures. *J. Med. Virol.* 92, 568–576. doi: 10.1002/jmv.25748
- WHO (2020). *World Health Organization Director-General’s opening remarks at the media briefing on COVID-19*. Geneva: WHO.
- Wong, H. Y. F., Lam, H. Y. S., Fong, A. H.-T., Leung, S. T., Chin, T. W.-Y., Lo, C. S. Y., et al. (2020). Frequency and distribution of chest radiographic findings in COVID-19 positive patients. *Radiology* 2020:201160. doi: 10.1148/radiol.2020201160
- Wright, A. F., Carothers, A. D., and Campbell, H. (2002). Gene–environment interactions—the BioBank UK study. *Pharmacogenom. J.* 2, 75–82. doi: 10.1038/sj.tpj.6500085
- Wu, J., Zhang, P., Zhang, L., Meng, W., Li, J., Tong, C., et al. (2020a). Rapid and accurate identification of COVID-19 infection through machine learning based on clinical available blood test results. *MedRxiv*. doi: 10.1101/2020.04.02.20051136
- Wu, J. T., Leung, K., and Leung, G. M. (2020b). Nowcasting and forecasting the potential domestic and international spread of the 2019-nCoV outbreak originating in Wuhan, China: a modelling study. *Lancet.* 395, 689–697. doi: 10.1016/S0140-6736(20)30260-9
- Wynants, L., Van Calster, B., Collins, G. S., Riley, R. D., Heinze, G., Schuit, E., et al. (2020). Prediction models for diagnosis and prognosis of covid-19: systematic review and critical appraisal. *BMJ* 369:m1328. doi: 10.1136/bmj.m1328
- Xu, B., Gutierrez, B., Mekaru, S., Sewalk, K., Goodwin, L., Loskill, A., et al. (2020b). Epidemiological data from the COVID-19 outbreak, real-time case information. *Sci. Data* 7, 1–6. doi: 10.1038/s41597-020-0448-0
- Xu, J., Glicksberg, B. S., Su, C., Walker, P., Bian, J., and Wang, F. (2020d). Federated learning for healthcare informatics. *J. Healthc. Informat. Res.* 2020, 1–19. doi: 10.1007/s41666-020-00082-4
- Xu, X., Jiang, X., Ma, C., Du, P., Li, X., Lv, S., et al. (2020c). A deep learning system to screen novel coronavirus disease 2019 pneumonia. *Engineering* 6, 1122–1129. doi: 10.1016/j.eng.2020.04.010
- Xu, Y., Li, X., Zhu, B., Liang, H., Fang, C., Gong, Y., et al. (2020a). Characteristics of pediatric SARS-CoV-2 infection and potential evidence for

- persistent fecal viral shedding. *Nat. Med.* 26, 502–505. doi: 10.1038/s41591-020-0817-4
- Yan, L., Zhang H-T, Goncalves, J., Xiao, Y., Wang, M., Guo, Y., et al. (2020b). An interpretable mortality prediction model for COVID-19 patients. *Nat. Mach. Intell.* 2020, 1–6. doi: 10.1038/s42256-020-0180-7
- Yan, L., Zhang, H.-T., Xiao, Y., Wang, M., Sun, C., Liang, J., et al. (2020a). Prediction of criticality in patients with severe Covid-19 infection using three clinical features: a machine learning-based prognostic model with clinical data in Wuhan. *MedRxiv*. doi: 10.1101/2020.02.27.20028027
- Yi, X., Walia, E., and Babyn, P. (2019). Generative adversarial network in medical imaging: a review. *Med. Image Anal.* 58:101552. doi: 10.1016/j.media.2019.101552
- Zaidan, A. A., Zaidan, B. B., Alsalem, M. A., Albahri, O. S., Albahri, A. S., and Qahtan, M. Y. (2020). Multi-agent learning neural network and Bayesian model for real-time IoT skin detectors: a new evaluation and benchmarking methodology. *Neural Comput. Appl.* 32, 8315–8366. doi: 10.1007/s00521-019-04325-3
- Zeng, H., Xu, C., Fan, J., Tang, Y., Deng, Q., Zhang, W., et al. (2020). Antibodies in infants born to mothers with COVID-19 pneumonia. *JAMA* 323, 1848–1849. doi: 10.1001/jama.2020.4861
- Zhang, J., Xie, Y., Pang, G., Liao, Z., Verjans, J., Li, W., et al. (2021). Viral pneumonia screening on chest X-rays using confidence-aware anomaly detection. *IEEE Trans Med Imaging* 40, 879–890. doi: 10.1109/TMI.2020.3040950
- Zhao, J., Zhang, Y., He, X., and Xie, P. (2020a). COVID-CT-Dataset: a CT scan dataset about COVID-19. *ArXiv [Preprint]*. *ArXiv: 200313865*. Available online at: <https://arxiv.org/abs/2003.13865v1>
- Zhao, S., Musa, S. S., Lin, Q., Ran, J., Yang, G., Wang, W., et al. (2020b). Estimating the unreported number of novel coronavirus (2019-nCoV) cases in China in the First Half of January 2020: a data-driven modelling analysis of the early outbreak. *J. Clin. Med.* 9:20388. doi: 10.3390/jcm9020388
- Zhou, F., Yu, T., Du, R., Fan, G., Liu, Y., Liu, Z., et al. (2020a). Clinical course and risk factors for mortality of adult inpatients with COVID-19 in Wuhan, China: a retrospective cohort study. *Lancet* 395, 1054–1062. doi: 10.1016/S0140-6736(20)30566-3
- Zhou, Y., Wang, F., Tang, J., Nussinov, R., and Cheng, F. (2020b). Artificial intelligence in COVID-19 drug repurposing. *Lancet Digit. Heal.* 2, 667–676. doi: 10.1016/S2589-7500(20)30192-8
- Zhu, X. J. (2005). *Semi-Supervised Learning Literature Survey*. Technical Report, University of Wisconsin, Madison, WI, United States.
- Zimmerman, A., and Kalra, D. (2020). Usefulness of machine learning in COVID-19 for the detection and prognosis of cardiovascular complications. *Rev. Cardiovasc. Med.* 21, 345–352. doi: 10.31083/j.rcm.2020.03.120

**Conflict of Interest:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2021 Abdulkareem and Petersen. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.