

The Role and Utility of Big Data and Artificial Intelligence in the Coronavirus (COVID-19) Pandemic Tracking and Containment

Siamak Zadeh*

Associate Professor and Data Scientist, Director, UG Program in Data Analytics, Ageno School of Business, Golden Gate University, San Francisco, California, USA

***Corresponding Author:** Siamak Zadeh, Associate Professor and Data Scientist, Director, UG Program in Data Analytics, Ageno School of Business, Golden Gate University, San Francisco, California, USA.

Received: April 09, 2020

Published: May 11, 2020

© All rights are reserved by **Siamak Zadeh**.

Abstract

This article aims to review the role and utility of big data and artificial intelligence in tracking and managing the public health crisis associated with the novel coronavirus (COVID-19) pandemic. The article reviews previous use cases where big data and AI have played a role and added value to monitoring and managing public health crises, outbreaks, and pandemics. The article discusses the benefits as well as the limitations and risks inherent in using big data and AI in managing public health crisis efforts. The article does not provide a critical review of the methodologies associated with the use of big data and AI for public health crisis management. It is hoped that the review provided here may facilitate and lead to future research to study and measure the efficacy of big data and AI in managing public health crises.

Keywords: Artificial Intelligence; Coronavirus (COVID-19); Pandemic

Introduction

The novel Coronavirus disease, COVID-19, has created an unprecedented global health crisis in the modern times. This outbreak started in December 2019 in Wuhan City of Hubei province, China [1,2] and within weeks spread globally. The World Health Organization (WHO) has officially declared this outbreak as a global pandemic [3,4]. At the time of this writing, this pandemic has infected millions of people worldwide and several hundred thousands have died from it in more than 160 countries (Figure 1) [5,6]. It is expected that these numbers will rise sharply before the pandemic is contained [6].

According to Centers for Disease Control and Prevention (CDC), SARS -CoV-2 is part of a family of zoonotic coronaviruses that have been detected in people and many different species of animals, such as camels, cattle, cats, and bats [5]. In the past, only a few types of respiratory diseases caused by coronaviruses, such as Middle East Respiratory Syndrome (MERS-CoV), Severe Acute Respiratory Syndrome (SARS-CoV), and now COVID-19, have infected people in great numbers and spread widely causing major pandemics and public health crises [7].

Preliminary studies and analyses of COVID-19 outbreak indicate that before this pandemic is fully understood and contained, millions more may be infected and hundreds of thousands may die

Figure 1: Confirmed COVID-19 cases globally.

Source: WHO Situation Report, March 17, 2020:
<https://www.who.int/emergencies/diseases/novel-coronavirus-2019/situation-reports/>

from this disease worldwide during the current pandemic period [7]. In addition, other early reports have indicated that the increase in the rate of infection will inevitably lead to severe shortages of critical health care resources, such as ventilators and Intensive Care Unit (ICU) capacity [8].

The current numbers of COVID-19 infected patients and deaths in the United States are also accelerating. Recent reports indicate

that the U.S. is experiencing increased community spread [9]. According to the U.S. federal government plan, the pace of the coronavirus pandemic is accelerating and may last at least 18 months, leading to increased infection and mortality rates as well as severe shortages of critical public health resources and capacity [10].

Big data and health care

With the continued growth in the volume and variety of global health data, researchers have studied the role and utility of big data to manage public health crises in general, and infectious diseases, outbreaks, and pandemics, in particular.

According to Dzau, *et al.* the amount of health-related data that can be captured, aggregated, stored, and analyzed continues to grow exponentially [11]. As a result, the opportunities to use the concept of the big data in all aspects of health care are expanding as well. Moreover, the increasing number of analytic tools that can extract meaningful clinical insights about the global healthcare situations from big datasets are leading to what Dzau has called a tipping point where big data will transform all aspects of health care operations, policies and practices [11,12].

For instance, Dzau predicted that the ever-growing volume and variety of data can provide insight into the state of health of the general public. Furthermore, it will allow various governments and the health care decision makers and administrators to identify problems early in their development and take actions in advance that could prevent some outbreaks from becoming major pandemics [12]. Clinical decision-support tools will query and mine data to recommend specific treatments tailored to specific situations, leading to a new era of efficient health care practices, as big data will be used to manage global health risks, from antimicrobial resistance to pandemic preparedness to public health risk stratification. It is this last area, according to Dzau that big data may make the biggest contribution to global health management, given that there are few risks to the world population that pose significant risks at the scale of infectious diseases and pandemics [13]. The need to access, collect and integrate infectious related data for public health crises management became very evident during the most recent outbreak of Ebola in West Africa, where the first case of Ebola was seen in December 2013 but for which WHO did not declare a global health emergency until August 2014 [11-13].

Big data ecosystem for infectious disease

The path to global health crises management requires access to large-scale, integrated, and detailed clinical and other health-related data to manage the onset, development, and spread of infectious diseases. Moreover, infectious disease management requires not only big data but detailed and diverse data. Advancing the field

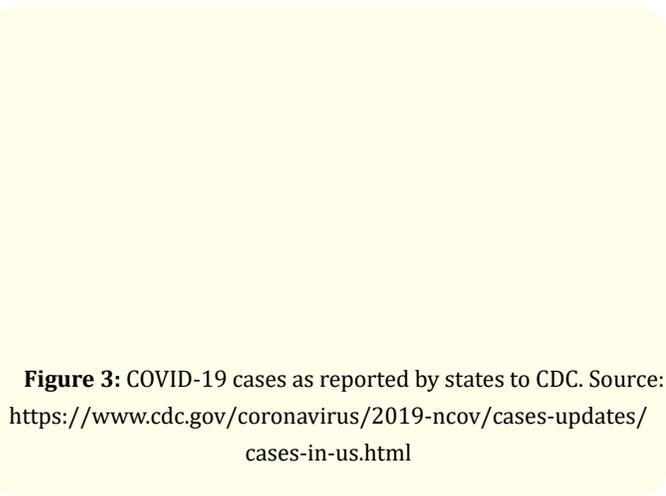
and improving understanding of the complexities of public health and disease will require integrating variety of datasets, structured, semi-structured, and unstructured, into a comprehensive clinical data repository [14]. Most of such repositories of research and clinical data are not interoperable and cannot be easily combined with one another.

With the growth in the adoption and utilization of electronic health records (EHRs) to collect, store and exchange data among healthcare providers, payers, and public health organizations and agencies, such as National institute of Health (NIH) and CDC, the infrastructure and the ecosystem for data-driven infectious disease management appear to be closer to the vision of a “health Internet”, an open platform that would connect providers, payers, and consumers to a growing number of applications, including ones designed for precision medicine [15].

One example of such an infrastructure and ecosystem is the CDC’s Surveillance Data Platform (SDP). SDP was designed to enable public health departments across the U.S. to send data related to injuries, diseases, and outbreaks to one place [16,17]. To facilitate the deployment of this infrastructure and allow researchers and technologists contribute to its development, the CDC launched the Health Information Innovation Consortium (CHIC) to ‘foster and promote creative solutions to surveillance challenges unique to public health’ [18]. Figure 2 depicts the strategic objectives and promise of the SDP initiative.

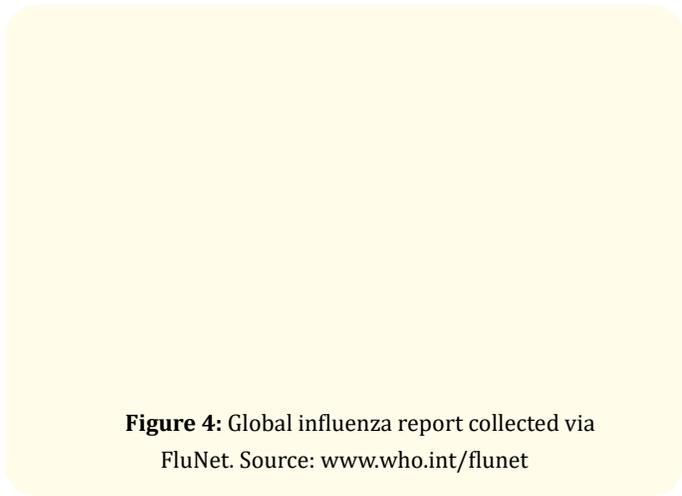
CDC is currently using this infrastructure to collect the number of COVID-19 cases across the U.S. (Figure 3).

Figure 2: Overview of the Surveillance Data Platform.
Source: CDC.



Similarly, WHO has established a platform for global reporting of influenza like infectious diseases. FluNet is a worldwide tool for influenza virological surveillance [19]. FluNet was first launched in 1997. All member countries report the virological data to FluNet, e.g. number of influenza viruses detected by subtype. These reports and numbers are critical for tracking the movement of viruses worldwide and interpreting the epidemiological data.

According to WHO, the data are made available by National Influenza Centers (NICs) of the Global Influenza Surveillance and Response System (GISRS) and other national influenza reference laboratories collaborating actively with GISRS or are uploaded from WHO regional databases [19]. Figure 4 shows an example of data collected by WHO through FluNet.



Such big data infrastructure and ecosystems are allowing the care providers and agencies responsible for the surveillance and monitoring to have information about the type, potency and transmission rates of infectious diseases, including the novel Coronavirus, COVID-19. Such access enables timely response to the impact-

ed areas and adoption of policies and regulations to mitigate the risk of such pandemics.

AI and healthcare

AI is increasingly being viewed as an important and supplemental tool, improving all aspects of health care provisioning, practices, population health management, and ultimately patient and clinical outcomes [20].

An early driver for use of AI in the health care industry stemmed from the comprehensive transformation that the U.S. health care industry is going through as it transitions from fee-for-service care model to a value-based and prevention-based approach. This transformation is being driven by the changes in the reimbursement policies and practices by both the public (i.e. Center for Medicare and Medicaid Services) and private health care payers (private insurers) [21-25]. This new approach aims to establish a wholistic and long-term view of a patient health care needs and provide a care management approach that reduces costs yet provides a better care overall.

In 2009, the Institute of Medicine (IOM) convened a roundtable to discuss the transition to the value-based care model and establish what has since been called 'evidence-based medicine' [26]. As with AI, evidence-based medicine requires access to fully integrated and detailed health care data as otherwise, the underlying characteristics of the process and outcomes are unknown. However, as outlined in the previous section, with the growth of the volume and variety of health care related data as well as the computational capacity throughout the health care systems, this approach and application of AI is becoming more feasible.

Table 1 highlights a few examples of how AI has thus had a positive impact on various aspects of health care services, procedures, and processes. As noted, AI has the promise to play a significant role in the management of wellness in general, and chronic diseases such as cardiovascular diseases, diabetes, and depression, in particular. These tasks can include a wide range of routine and repeated tasks such as taking medications, exercises, and vital signs to help assess the general physical health condition of a person [20]. AI algorithms can be developed and trained to review the data from such monitoring to assess the changes in a patient condition and provide alerts if anomalies are detected in the person's general patterns. Routine and repeated tasks are often good candidates for automation through AI.

AI is also transforming how clinical care is delivered by providers across the spectrum of care deliveries. According to the National Academy of Medicine, "there are two main areas of opportunity

Use Case	Category	Examples of Applications	Technology
Patients and Families	Health Monitoring	Devices and wearables	Machine Learning, Natural Language Processing (NLP), Speech recognition
	Disease Prevention and Management	Obesity reduction Diabetes prevention and management Emotional and mental health support	Conversational AI, NLP, Speech recognition, mood analysis
Clinicians and Care	Early Detection, prevention, and diagnostic tools	Early disease detection	Machine learning
Providers	Surgical procedures	AI-supported surgical roadmaps	Robotics, machine learning
	Precision medicine	Personalized chemotherapy	Supervised machine learning
	Patient Safety	Early detection of sepsis	Machine Learning

Table 1: Example of AI in health care.

Source: Adopted and modified from Chapter 3 [20].

for AI in clinical care: (1) enhancing and optimizing care delivery and (2) improving information management, user experience, and cognitive support in electronic health records (EHRs)” [20].

Another area that AI has had some early impacts is in radiology and pathology as imaging analytics is already helping clinicians with better detection and diagnostics. Computer-aided detection (CADe), also known as computer-aided diagnosis (CADx), are techniques that assist physicians in the interpretation of clinical images [27]. Digital images from X-ray, MRI, and ultrasound systems generate a vast amount of data that if analyzed properly can help with improved diagnosis and prognosis. Processing these digital images by CAD systems are often automated to assist clinicians and health care practitioners with diagnosis in a relatively short time. Moreover, AI-based systems have the promise to improve digital pathology with the advent of ‘whole-slide imaging’ and machine learning [28,29].

AI can also improve and streamline the administrative processes of health care. Automation and machine learning can help reduce the administrative gaps and reduce costs significantly [28].

AI and Infections Disease

The availability of large-scale data from the health-related big data ecosystems, such as those discussed above, has enabled epidemiologists, infectious disease monitoring agencies such as WHO and CDC, and policy makers to utilize AI and machine learning to study, track, and manage the onset, progression, and the spread of outbreaks. Some recent studies have shown early promises in tracking zoonotic infectious diseases from rodent reservoirs [30]. In another study, researchers at the Viterbi School of Engineering at the University of Southern California in the U.S. developed an algorithm to help policy makers reduce the spread of infectious diseases such as tuberculosis (TB), malaria, and gonorrhoea [31]. Utilizing AI-based simulations, the researchers applied the algorithm to two real-world cases: TB in India and gonorrhoea in the U.S. The study found the AI algorithm performed better at reducing disease cases than outreach programs by disseminating information about these diseases among vulnerable individuals [31].

Most applications of AI and machine learning to infectious disease tracking, monitoring, and containment generally follow the steps depicted in figure 5. As in the case of COVID-19, the first step is to identify the pathogen such as SAR-Cov-2 as in the case of COVID-19. Once the pathogen and the transmission dynamics of the infectious disease are identified, epidemiological and virological studies as well as population health data can be collected and aggregated to provide input into various AI algorithms to analyze the data and provide updated information and guidelines to the policy makers, clinicians, and the public at large. AI and machine learning can utilize such epidemiological studies and data to predict an epidemic from its onset through various stages, providing various scenarios for the policy makers to help them contain the spread of the disease [32,33]. The promise of such AI-based approaches is that the algorithms and the models can be updated regularly as additional information and data become available in real time.

A key requirement for the utility of AI and machine learning to successfully track, monitor, and help contain any pandemics is the development of viable algorithms to model and simulate the transmission dynamics of the infectious diseases. In the past both deterministic as well as stochastic models have been used to simulate such transmission dynamics [34-38]. Though these models have generally helped with better understanding and prediction of pandemics, the overall role and utility of AI and machine learning to accurately identify infectious disease spread and pandemics is in its early stage [39]. There are still challenges to accessing fully integrated and detailed epidemiological data across all disparate sources.

Figure 5: General flow of the application of AI to the study of infectious diseases from first identifying the pathogen, to the study of the transmission and spread of the disease, to data collection from individuals as well as care providers, to the final application of data analytics and artificial intelligence.

In an open letter, the Data, a membership-based trade association in the U.S. that advocates for responsible policies to make government data high-quality, accessible and useable, has called on the U.S. congress to set \$50 million this year and \$125 million in fiscal 2021 to support the rollout of the Foundations for Evidence-Based Policymaking Act, in response to COVID-19 pandemic [40]. In this letter, the coalition stresses upon access to a 'high-quality data and a robust data infrastructure' to address the current as well as future pandemics to better prepare for future public health crises.

Concluding Remarks

The promise of big data and AI in health care in general, and fighting pandemics in particular, offers significant opportunities to address public health crises. These tools can provide timely data analysis to facilitate tracking and monitoring of the spread of infectious diseases and help clinicians, epidemiologists, agencies such as WHO, CDC, and other public health entities to better understand the transmission dynamics of the infectious diseases and provide more effective guidelines and policies during a pandemic. A good example of this is the simulation and modeling of the COVID-19 pandemic by CDC to represent the impact of two different social approaches, shelter-in or no shelter-in guidelines (Figure 6).

Big data and AI are already playing a major role in transforming a variety of businesses. Despite its nascent stage, the benefits of application of AI and machine learning to health care and containment of pandemics outweighs any limitations. Improving the national health care data infrastructure and providing additional computational resources will inevitably help in better data analysis and development of more accurate predictive models. AI needs to

Figure 6: Simulation and modeling of two different approaches, such as no social distancing vs social distancing by public, on the spread of COVID-19. Source: CDC.

become an integral and supplementary part of health care management, particularly to fight the onset and the spread of infectious diseases.

Bibliography

1. Juanjuan Zhang, *et al.* "Evolving epidemiology of novel coronavirus diseases 2019 and possible interruption of local transmission outside Hubei Province in China: a descriptive and modeling study". (2020).
2. Zifeng Yang, *et al.* "Modified SEIR and AI prediction of the epidemics trend of COVID-19 in China under public health interventions". *Journal of Thoracic Disease* (2020).
3. World Health Organization. WHO Director-General's opening remarks at the media briefing on COVID-19 - 11 March 2020 (2020).
4. World Health Organization Declares COVID-19 a 'Pandemic.' Here's What That Means.
5. World Health Organization COVID-19 Dashboard (2020).
6. John Hopkins University Center for Systems Science and Engineering COVID-19 Dashboard (2020).
7. Centers for Disease Control and Prevention (CDC) (2020).
8. Giacomo Grasselli, *et al.* "Critical Care Utilization for the COVID-19 Outbreak Published in Lonlinem bardy, March Italy 13, Early2020 Experience. and Forecast During an Emergency Response". *JAMA* (2020).
9. Centers for Disease Control and Prevention (CDC) (2020).
10. US plan warns coronavirus pandemic could last 18 months: report.

11. Dzau VJ., *et al.* "Realizing the Full Potential of Precision Medicine in Health and Health Care: A Vital Direction for Health and Health Care". NAM Perspectives. Discussion Paper, National Academy of Medicine, Washington, DC (2016).
12. National Academies of Sciences, Engineering, and Medicine. Big data and analytics for infectious disease research, operations, and policy: Proceedings of a workshop. Washington, DC: The National Academies Press (2016).
13. Dzau VJ., *et al.* "Aligning incentives to fulfill the promise of personalised medicine". *Lancet* 385.9982 (2015): 2118-2119.
14. NRC (National Research Council). Toward precision medicine: Building a knowledge network for biomedical research and a new taxonomy of disease. Washington, DC: The National Academies Press (2011).
15. Kapor M. "Building the Health Internet. Presented at The Meeting at Harvard on a Health Information Technology Platform, Boston, MA (2009).
16. Center for Disease Control and Prevention: Public Health Surveillance Preparing For The Future. (2020).
17. Center for Disease Control and Prevention: Public Health Data Modernization Initiative: Harnessing The Power Of Data To Save Lives (2020).
18. Center for Disease Control and Prevention: Surveillance Strategy Report - Surveillance Innovation (2020).
19. World Health Organization: Flu Net.
20. Matheny MS., *et al.* "Artificial Intelligence in Health Care: The Hope, the Hype, the Promise, the Peril". NAM Special Publication. Washington, DC: National Academy of Medicine (2019).
21. Alley DE. "Accountable Health Communities - Addressing Social Needs through Medicare and Medicaid". *New England Journal of Medicine* 374 (2016): 8-11.
22. ASTHO (Association of State and Territorial Health Officials). Medicaid and Public Health Partnership Learning Series (2019).
23. CMS (Center for Medicare and Medicaid Services). Medicare and Medicaid Programs; Patient Protection and Affordable Care Act; Interoperability and Patient Access for Medicare Advantage Organization and Medicaid Managed Care Plans, State Medicaid Agencies, CHIP Agencies and CHIP Managed Care Entities, Issuers of Qualified Health Plans in the Federally-Facilitated Exchange and Health Care Providers. Proposed Rules. Federal Register 84.42 (2019): 7610-7680.
24. Kissam SM., *et al.* "States encouraging value-based payment: Lessons from CMS's state innovation models initiative". *The Milbank Quarterly* 97.2 (2019): 506- 542.
25. Mendelson A., *et al.* "The effects of pay-for performance programs on health, healthcare use, and processes of care: A systematic review". *Annals of Internal Medicine* 166.5 (2017): 341-355.
26. Institute of Medicine (US) Roundtable on Evidence-Based Medicine. Leadership Commitments to Improve Value in Healthcare: Finding Common Ground: Workshop Summary. Washington (DC): National Academies Press (US) (2009).
27. Computer-aided diagnosis.
28. Bresnick Jennifer. Identifying Near-Term Use Cases for AI in the Healthcare Setting (2019).
29. Bresnick Jennifer. Artificial Intelligence Promises a New Paradigm for Healthcare (2019).
30. Han BA., *et al.* "Rodent reservoirs of future zoonotic diseases". *Proceedings of the National Academy of Sciences of the United States of America* 112.22 (2015): 7039-7044.
31. Blumenthal A. "Artificial Intelligence to fight the spread of infectious diseases". Phys. Org. (2018).
32. Agrebi S and Larbi A. "Use of artificial intelligence in infectious diseases". In Artificial Intelligence in Precision Health; From concept to applications. Edition. Barth, D. Elsevier Inc (2020): 415-438.
33. Majumdar A., *et al.* "Kysanur Forest Disease Classification Framework Using Novel Extremal Optimization Tuned Neural Network in Fog Computing Environment". *The Journal of Medical Systems* 42 (2018): 187.
34. Longini IM., *et al.* "Containing Pandemic Influenza at the Source". *Science* 309.5737 (2005): 1083-1087.
35. Sander B., *et al.* "Economic Evaluation of Influenza Pandemic Mitigation Strategies in the United States Using a Stochastic Microsimulation Transmission Model". *Value in Health* 12.2 (2009): 226-233.
36. Van den Dool C., *et al.* "Modeling the Effects of Influenza Vaccination of Health Care Workers in Hospital Departments". *Vaccine* 27.44 (2007): 6261-6267.
37. Lizon NE., *et al.* "Incorporating Healthcare Systems in Pandemic Models". Proc Winter Simulation Conf (2010): 2230-2236.
38. Andradóttir S., *et al.* "Reactive Strategies for Containing Developing Outbreaks of Pandemic Influenza". *BMC Public Health* 11.1 (2011).

39. Tsui KL, *et al.* "Tracking infectious disease spread for global pandemic containment". *IEEE Intelligence Systems* 28.6 (2013): 60-64.
40. An Open Letter to Congress for Responding to the COVID-19 Pandemic with Data (2020).

Assets from publication with us

- Prompt Acknowledgement after receiving the article
- Thorough Double blinded peer review
- Rapid Publication
- Issue of Publication Certificate
- High visibility of your Published work

Website: www.actascientific.com/

Submit Article: www.actascientific.com/submission.php

Email us: editor@actascientific.com

Contact us: +91 9182824667