Chapter 9
Effectiveness of Big Data in Early Prediction and Measure for COVID-19 Using Data Science

P. Tomar, M. Mann, D. Panwar, C. Diwaker, and P. Kumar

Abstract  COVID-19 originates from a group of infections related to severe acute respiratory syndrome (SARS). Through data analysis, frameworks have demonstrated to have the option to approximate how patterns may progress. Data analysis assumes a key role, as does arithmetic, which, together with data science, permits us to have a top to bottom comprehension of the subtleties of nature and how things are made. As in the previous years, the pioneers of data science have had a mind-blowing effect on the reality where information and examination have been utilized to drive critical change throughout the spread of the sickness. One of the primary recorded uses of information examination was in 1852, during a cholera episode in London. John Snow, one of the primary information-driven disease transmission specialists, had the option to geospatially examine the passing that happened in London and in this manner segregate the wellspring of the infection. Depending on his investigation, specialists had the option to focus on their interventions to quickly check the spread of the pandemic. During this pandemic, information can be an exceptional factor as far as quality and consistency. Complexities of this sort incorporate instances of false-positive patients. Large data and data science can be utilized to check consistency with isolate and can be utilized for tranquilizing
investigation. These are only a portion of the arrangements offered by new advanced innovations in the field of data science to confront the coronavirus crisis. We provide here early prediction and early survey, with the genuine and potential commitment of data science to the battle against COVID-19, just as the present requirements on these commitments. It expects to draw brisk take-away from a quick extending conversation and developing a collection of work, to fill in as a contribution for fast reactions in research, approach, and clinical examination. The expense of the pandemic as far as lives and monetary harm will be horrible; at the hour of composing, incredible vulnerability encompassed appraisals of exactly how awful and of how fruitful both non-pharmaceutical and pharmaceutical reactions can be. Improving data science, one of the most encouraging information expository apparatuses to have been created over the previous decade or something like that, to help decrease these vulnerabilities, is a beneficial interest. Data science isn’t yet assuming a huge job in the battle against COVID-19, at any rate from the epidemiological, demonstrative, and pharmaceutical perspectives. Its utilization is compelled by an absence of data and by an excess of noisy data. Enthusiastically, data analysts and researchers have responded to the challenge.

**Keywords** Deep learning · Breast Cancer Coimbra · Immunotherapy · Attribute selection

### 9.1 Introduction

COVID-19 originates from a group of infections related to severe acute respiratory syndrome (SARS) [1]. Coronavirus illness (COVID-19) is an aggravation infection from another infection. The illness causes respiratory affliction (like flu) with signs, for instance, cool, hack, and fever, and, in dynamically genuine cases, the issue in relaxing. COVID-19 has been seen as an overall pandemic [2].

As of date affirmed COVID-19 cases across the globe are 1,498,833 and mortality roughly 5.8%. Continuously the death rate is expanding and it’s a disturbing component for the entire world. Transmission is ordered into four phases depending on the method of spread and time [3]. Each country forced various systems beginning from remaining in-home, utilizing covers, travel limitations, staying away from parties, regularly washing hands, and cleaning the spots frequently on the account of a typical exertion to battle the flare-up of this ailment. Numerous nations forced a lockdown express that forestalls the development of the residents pointlessly [4].

In India, the flare-up of coronavirus has upset the working of life in general. All were pushed to remain back to the shield from the repulsive transmission. In the
underlying stages, the affirmed cases are those that come back from administers followed by transmission using nearby transmission [5].

“Huge information” is a monstrous measure of data that can be unbelievable. For over two decades, it has become a subject of extraordinary enthusiasm given its incredible potential [6]. Various transparent and private corporations produce, store, and study immense knowledge to strengthen their administrations. In the healthcare sector, administrative reports, hospital clinical data, and implications of psychiatric tests and online apps provide various hotspots with massive volumes of knowledge [7]. A significant part of broad knowledge related to open healthcare is also derived from the biomedical examination [8]. The legitimate administration and scrutiny of this information are needed to infer important data. Someone else that is almost identical to finding a needle in the bundle by dissecting enormous information fast. There are various difficulties in each step of the management of extensive information that must be executed by using top-of-the-line registration solutions for an enormous analysis of information [9]. That is why providers of medical care need to be equipped entirely with the correct structure to build and break down tremendous knowledge in a methodical way and provide practical responses and enhance general health.

9.2 Background

There are several promising and life-saving outcomes in the use of broad knowledge analysis in social security. Widespread information refers to the huge amount of data produced by the digitalization of all that is united and explicitly breached by innovations. It uses specific wellbeing details from the community (or from a single person) to assist in the prevention of plagues, fixing illness, capping expenditures, etc. [10], as used in public resources.

Since we live longer, models of treatment have changed, and many of these advances must be driven by information. Specialists must be as much as possible aware of a patient and, from the start, as reasonably expected, to obtain noticed signs of real disorder when emerging—it is unmistakably basic and more affordable to reward any disease at first. In the analysis of knowledge for healthcare, avoidance is preferable to deciding and whether a patient should be identified in a wide-ranging manner to provide security a personalized package. The corporation seeks to resolve the difficulties the knowledge of a patient contains: everywhere it is processed and put in emergency departments, hospitals, surgical treatments, etc., with the inability to interact properly [11].

It has been exorbitant and tedious for quite a long time to collect enormous information measures for clinical use. With the progress made at present, it is simpler not only to gather this information but also to transform it into important basic pieces of knowledge, which could be used to better reflect. The reasons behind the review of information for human services are using information-led findings to anticipate and deal with a problem before the no return is over, but also survey policies and
medicines more quickly, monitor stock, incorporate patients further into their welfare, and enable them to take care of themselves.

9.2.1 Big Data Analytics in Healthcare

The need for massive information is also colossal in medicinal services, as expenses in countries such as the United States are increasing. As a survey by McKinsey says, “After over 20 years of consistent builds, medicinal services costs currently speak to 17.6 percent of GDP—almost $600 billion more than the normal benchmark for a country of the United States’ size and riches” [12].

Costs are far larger, so to say, than they should be, and for as much as 20 years, they have grown. Here we simply need a shrewd, insightful deduction. Also, the driving forces of today change: Many insurance companies are shifting from administrative charges to schemes that organize lasting results (which use expensive, now and again usually useless medicines and recompensing many patients quickly) [12].

In the past, all the considerations showed no immediate impetus for the suppliers of medicinal services to provide persistent, intensive use of research data to each other. As more get charged based on the interpretation of outcomes, there is a cash-related opportunity for information to be exchanged that can be used to enhance patients’ lives and lower costs for defensive companies [13].

Finally, doctoral options have become more and more validated, meaning that they are not solely dependent on tutoring and professional appraisal, but on broad scopes of research and clinical knowledge. As with numerous companies, social information events and the board are growing, and experts need help. This latest research means that there is a greater interest than ever in recent memory in broad knowledge review at medical facilities, and the emergence of SaaS BI devices further suggests that they need to be tested [9].

9.2.2 Effectiveness of Big Data Against COVID-19

Near to COVID-19, which constantly gather data from sites across the world, social security staff, scholars, experts in health communication, and decision-makers are supporting global episode information.

There has been some fascinating information coming about because of GPS investigations of populace development by district, city, and so on, which at last gives a perspective on the populace’s consistency—or absence of consistency—with social removing mandates. Moreover, in any event, an advanced thermometer invention follows temperatures reported by their devices, suggesting that, despite approval of a social repression regulation, normal temperatures in some areas have declined [14].
How have advances made it possible to source, aggregate, and synthesize incident data globally in Big Data management and data preparation?

For such cases like these, there are various chances of using Big Data as a general population and an enterprise, while the strength of Big Data for the search of a solution has still not been made viable [15].

For example, considerations of large-scope COVID-19 real-world evidence (RWE) could help to boost medicines’ progress in a progressive and patient-friendly manner, drawing information from several true sources—including patient care currently being treated in medical clinic environments [16].

The pandemic of COVID-19 could be seen as an inspiration for determining how to improve access to information. We at Paraxial try to ally with other experts in medication improvement and industry relations to help with this exercise in the United States, but we must move quickly [17].

What can Big Data allow a more data-driven and proactive battle against global pandemics?

As both the general public and industry see more noteworthy progress, we still have to take measures to make viable use of large data intensity in the search for a solution to COVID-19. Progressive examination and sign findings within the framework of social insurance are one of some huge improvements in the computerization of the early pandemic surface. There will be more variations in a regional context. Large data models improved pandemic standards. Our ability to absorb and respond to therapy depends early and on our efforts to use as much exposure knowledge as possible. This will help us to understand contrasts in the introduction, to respond to the various therapeutic modalities, and to codify the drugs that have a better impact.

How does Big Data contribute to clinical trials’ efficiency and effectiveness?

During this time, clinical preliminary tasks take full advantage of the information’s capacity to interface us with the patients. The result is a faster, more focused access to these significant studies and patient entry while ensuring quiet access [18].

We at Paraxial are investigating approaches to make huge-scope COVID-19 RWE considerations that can pull information from what we are seeing with patients right now under treatment. A portion of the potential advantages incorporates an increasingly understanding-driven and patient-accommodating spotlight on recognizing and creating potential cures [19].

The key advantage is an open door for quicker and close to the constant assessment of dynamic dependent on the information. This will spare lives and result in recognizing viable treatments quicker.

Another thought is to distinguish expected patients at the purpose of testing and afterward quickly populate understanding narratives and track progress by connecting to different genuine sources using tokenization with constrained weight on the patient and the site.

When medicinal services information can be utilized straightforwardly (e.g., engineered control arms) to construct the clinical informational collection for well-being and adequacy examination, fewer patients should be selected and
randomized, lessening the complete span of the investigation and the calculated weight on patients.

9.3 The Outbreak of COVID-19 and Exponential Growth of COVID-19 Cases

The ongoing flare-up of coronavirus sickness 2019 (COVID-19) in territory China was described by an unmistakable subexponential increment of affirmed cases during the early period of the plague, appearing differently about an underlying exponential development expected for an unconstrained episode. We show that this impact can be clarified as an immediate result of control approaches that adequately exhaust the powerless population [20].

On the account of COVID-19, exponential development will happen in the ailment rate in people in as much as:

- There is at any rate one tainted individual in the populace pool
- Regular contact among contaminated and uninfected individuals from the populace happens
- There are enormous quantities of uninfected likely among the populace

Coronavirus spreads exponentially instead of straightly. Exponential development inclination makes it difficult for residents to assess the status of the pandemic in their nations remembering the expected results for inevitable days [21]. A significant point in our COVID-19 arrangement of blog entries is to introduce and give intelligent representations and instruments, understandable by the non-master, that will assist in facilitating the exponential development predisposition obstruction and shed some light on patterns and relationships supporting COVID-19 data [21].

When you have a progression of information from date A to date B, we are not fit for imaging the information to come during the next days. All things considered, here comes the guide of prescient displaying! We can take a model that is compelled to comply with exponential development conduct and tune it in an information-driven way (e.g., utilizing the accessible information from date A to date B) to fit them best. This activity is deciphered as follows: among all conceivable exponential development bends (which are interminable), to discover and utilize the one that follows the accessible information as firmly as could reasonably be expected.

Government restrictions have been actively carried out in China on the production of corona viral diseases 2019 (COVID-19) as a result of severe extreme respiratory coronavirus 2 (SARS-CoV-2). The fast increment of emphatically analyzed cases and a resulting ascent of auxiliary flare-ups in numerous nations overall raised worry on a universal scale. The World Health Organization (WHO) in this way reported the COVID-19 episode The Global Public Health Crisis on 31 January, which was eventually declared as a pandemic on 11 March [22].
An underlying exponential development of affirmed cases is conventionally expected for an uncontrolled flare-up, as watched, e.g., during the 2009 flu A (H1N1) pandemic or the 2014 Ebola episode in West Africa [2]. Even though in Hubei the quantity of research facility-affirmed cases $C(t)$ was seen to develop exponentially toward the beginning of January, the ensuing ascent followed a sub-exponential, superlinear, mathematical scaling law $t\mu$ with a type $\mu = 2.3$ (between 24 January and 9 February). For the greater part of the influenced Chinese territories of terrain China, nonetheless, this sort of mathematical ascent happened from the earliest starting point of the case giving an account of 21 January [23]. The examples $\mu$ change around a run of the mill estimation of $\mu = 2.1 \pm 0.3$ for the affirmed case bends in other generously influenced territories (affirmed case tallies bigger than 500 on 12 February), showing mathematical development notwithstanding land, socioeconomic contrasts, potential contrasts in control systems, and heterogeneities that may impact and affect how the neighborhood pestilence unfurls. In the end, case checks started to go amiss from the watched scaling laws around 9 February for Hubei and toward the beginning of February for the rest of the territories, moving toward an immersion behavior [23].

9.4 Predicting Survival Chances Using Data Sciences

9.4.1 Finding Who Is Most at Danger from COVID-19

AI has demonstrated to be significant in anticipating dangers in numerous circles. With clinical hazard explicitly, AI is conceivably fascinating in three key manners.

- Infection chance: What is the danger of a particular individual or gathering getting COVID-19?
- COVID-19 Chapter International Disaster Public Health Warning on 31 January and eventually on 11 March [22] as a pandemic.
- Outcome hazard: What is the hazard that a particular treatment will be incapable for someone in particular or gathering, and how likely would they say they are to bite the dust?

AI can help foresee every one of the three dangers. Also, we can see how AI is used in associated areas and how it could contribute to the COVID-19 chance forecast [24].
9.4.2 Infection Risk Estimation

Early statistics indicate that important risk factors that decide how likely an individual is to get COVID-19 are age, pre-existing conditions, general hygiene habits, social habits, number of human interactions, frequency of interactions, location and climate, and socioeconomic status.

In the early stages, there is still potential analysis for the pandemic ebb and flow. For example, DeCapprio et al. used AI to create an underlying COVID-19 vulnerability index. Avoidance estimates, for example, wearing veils, washing hands, and social separating, are for the most part prone to impact generally speaking danger also. As more and better information opens up and as of now progressing investigations produce results, we will probably observe progressively pragmatic uses of AI for foreseeing contamination risk [25].

9.4.3 Assessment of Who Is at Risk of a Critical Case

If a person or gathering has been contaminated, we need to foresee the risk of that person or collect problems or require propelled clinical consideration. Some people experience only minor side effects, and other individuals cause severe pulmonary disease or serious respiratory disorder (ARDS) that is considered to be harmful. Nonetheless, it is better to begin therapy early if more serious symptoms are expected to occur, as they all need to be properly handled and screened for soft signs [26].

In the Computers, Materials and Continua diary, specialists distributed an article demonstrating that AI might anticipate the probability of a patient creating ARDS just as the danger of mortality, just by taking a gander at the underlying side effects. The specialists recognize the restrictions of this examination:

“An away from this examination is the size of the dataset; 53 patients with some inadequate information just as a restricted range of seriousness.”

9.4.4 Predicting Treatment Outcomes

An augmentation of seriousness forecast is anticipating the treatment’s result, which is frequently truly a matter of foreseeing life and demise. It is valuable to realize how likely a patient is to endure, given certain side effects. However, on this present, it’s critical to remember that not all patients are treated similarly. Given a particular patient or gathering, how compelling is a particular treatment liable to be?
If we can anticipate the results of explicit treatment strategies, at that point specialists can treat patients all the more adequately. Utilizing AI to customize treatment plans isn’t explicit to COVID-19, and AI has recently been utilized to anticipate treatment results for patients with epilepsy, as only one model. Scientists have likewise utilized AI to foresee reactions to malignant growth immunotherapy [16].

Since treatment alternatives for COVID-19 are as yet developing, it will probably be some time before we see AI applied to foresee results for explicit medicines. Be that as it may, result expectation stays a significant piece of hazard evaluation, working connected at the hip with the disease and seriousness forecasts we talked about above.

9.5 Fighting COVID-19 Using Big Data and Data Science

COVID-19 has transformed us, how we communicate, our propensities, and even our psyches. It is genuine that we are battling the infection from numerous points of view. Social removing, immunizations, and clinical medicines are the most notable, yet additionally, numerous different methodologies have risen, some of them from the data science and AI people group [27].

One of them is the COVID-19 Mobility Monitoring venture in Italy, which utilizes information gathered from anonymized clients who have been selected to give access to their area information namelessly, through a GDPR-agreeable system. Information is gathered from cell phones with iOS and android utilizing various sensors, for example, GPS, signals, Wi-Fi, and systems. They will probably quantitatively investigate the effect of the imprisonment and social removing measures forced in Italy [28].

Information science is an entirely important instrument that researchers and governments are utilizing to adapt to the impacts of the COVID-19 pandemic, and numerous incredible applications are rising at an extremely quick pace. It is exceptionally compensating to realize that our field is assisting with this worldwide emergency and it can help with numerous different issues of our advanced society [28].

9.5.1 Role of Data Science in Global Health Emergency

The tremendous knowledge strategy should talk of a major development in the area of human resources outreach—promoting the improvement of the instructional frames in treatment and ultimately promoting effective control of individuals to enhance the health of whole communities. The size of the approach increases both the expected dangers and potential benefits. Although the methodology can have the greatest incentive for low-asset settings, it is generally powerless in such settings against discontinuity and abuse. The vulnerability is supposed to be restricted by community-oriented administration, diligent analysis, and specialist organizations.
The dynamics should not be taken for granted. Another opportunity for worldwide wellbeing institutions to provide valuable governance in countries that are low-and-center salaries is the shift from paper to petabyte computerized supplies [29].

A working social insurance framework is essential for the endurance of mankind. Without specialists, medicines, and malady counteraction measures, we’d return to the Dark Ages (and still, after all that, there were specialists). Modern social insurance gear produces a great deal of wellbeing information, and this is the place Big Data applications can help [30].

### 9.5.2 Virus and Drug Research Using Data Science

It costs up to $2.6 billion and takes 12 years to put up medication for sale to the public. Large information permits researchers to reproduce the response of medication with body proteins and various kinds of cells and conditions, so it has a lot higher probability of picking up the Food and Drug Administration endorsement and restoring assorted patients (e.g., individuals with certain change profiles) [31].

Informatics can use different arrangements for organized and unstructured biomedical information obtained from diverse tests, therapeutic results, contextual studies, online life, etc. Such advanced scientific calculations would then be used to recreate how the medicines will work with body proteins and how the pace of success is anticipated [32].

The recreation can accelerate the procedure making beginning screening adequate to decide the chance of the medication viability. In addition to the fact that it means a tremendous decrease in the expense and time of medication advancement, however, it additionally mitigates the dangers of disappointment.

Information science strategies are likewise ready to incorporate with genomic information in exploration to give exact bits of knowledge into hereditary issues emerging out of explicit medications and sicknesses.

Medication improvement is a tedious and complex excursion, with a high vulnerability that medication will succeed, while the rise of a major information approach has altered our methodologies to handle long-standing difficulties in tranquilizing advancement. The proposition of different high-throughput innovations and an assortment of numerous omics information are quickening the interpretation of essential exploration revelations into clinical practice. The previous decade has seen the interpretation of a few examination results from genomics information to FDA endorsement in facilities, and in the interim, numerous as of late created information-driven methodologies have additionally demonstrated promising potential for clinical practice. An early and surely understand model is headache medicine, which is initially utilized for the absence of pain treatment, while by coordinating the data from the electronic wellbeing records (EHRs) of patients, post-advertising reconnaissance information, and pharmacological investigation, specialists likewise found the capability of anti-inflammatory medicine to treat colorectal disease, and
the US Preventive Services Task Force discharged the draft suggestion on ibuprofen to forestall colorectal malignancy in September 2015 [33].

9.5.3 Analysis and Prediction of Foreign Epidemic in the Future

Formal, quantitative methodologies are presently broadly used to make expectations about the probability of an irresistible infection flare-up, how the ailment will spread, and how to control it. A few entrenched techniques are accessible, including hazard factor investigation, chance demonstrating, and dynamic modeling [34]. All things considered, prescient demonstrating is particularly the “specialty of the conceivable,” which will in general drive research exertion toward certain regions and away from others which might be at any rate as significant. Expanding on the undoubted achievement of quantitative demonstrating of the study of disease transmission and control of human and creature illnesses, for example, AIDS, flu, foot-and-mouth ailment, and BSE, consideration should be paid to building up an increasingly all-encompassing system that catches the job of the basic drivers of ailment dangers, from demography and conduct to land use and environmental change.

None of us can precisely and completely anticipate the results of any pandemic, the all-out disease rates, or the last loss of life as of right now. These are troublesome occasions for all of us, on the account of the concern and the vulnerability of the circumstance. We just don’t have the foggiest idea of what will occur, yet with information science, we have a superior possibility of foreseeing it precisely than we ever have had in the past [35].

Yet, even once the pandemic is finished, there will at present be information science work to do. The information assortment may stop because there are no new cases, yet that is an ideal opportunity to ensure that what was gathered is put away and overseen appropriately. Glancing back at the circumstance with the advantage of knowing the past can be agonizing, yet it’s the main way we can guarantee that exercises picked up during this time are completely comprehended. After everything closes is likewise an opportunity to take a gander at the frameworks and structures (clinical, logical, social) that worked, and the ones that didn’t, with the perspective on progress.

9.5.4 Future and Roles of Data Science in Prediction and Measurement for COVID-19

The precision of customary determining to a great extent relies upon the accessibility of information to base its forecasts and gauges of vulnerability. In episodes of pandemics, there is no information at all to start with and afterward constrained over the long haul, making forecasts broadly dubious.
Propelled representation has helped governments and analysts intently watch everyday improvements of COVID-19 and settle on choices successfully. Descriptive insights as well as examining the relationship among different variables are permitting chiefs to think about and comprehend the effect of the pandemic. Different associations are preparing a gigantic measure of information and giving perception to exhibit how the infection got out [36].

Information researchers are anticipating the spread of COVID-19 alongside data on the number of lives it is probably going to influence. Notwithstanding, anticipating with inadequate information can additionally befuddle individuals during these difficult occasions. Existing datasets (on COVID-19) are staggeringly one-sided. For example, while ascertaining the death rate, typically engineers take a gander at the passings per affirmed case. In any case, the supposition that will be that we have caught the entirety of the affirmed cases, which isn’t valid [37].

The job of information science is basic to mentioning information-driven objective facts on how the COVID-19 pandemic is spreading. About pandemic, clinical information can be profoundly factor regarding quality just as consistency; complexities of this incorporate instances of bogus positive patients. In any case, governments and specialists are depending progressively on information and forecast models in the plague setting [37].

Information researchers are utilizing huge information for AI calculations fit for creating a large number of restorative antibodies to discover medicines for COVID-19 with a high likelihood of progress. Leading immune response disclosure in a lab regularly takes years; in any case, these calculations recognize antibodies that can battle against the infection in only 7 days. This is only one of the numerous ways information-driven arrangements are quickening the speed at which coronavirus treatments are found, created, and delivered [37].

9.6 Conclusion

The spread of virus is exponential, and the application of data science in optimal measurements for outcomes of exponential model has been gratefully acknowledged in the recent past. This paper suggests how to fit COVID data as an exponential fit model. Different factors responsible for capturing this deadly virus are identified, and corresponding critical case scenarios are discussed in this study. The argumentation on predicting various available treatment and possible outcomes are studied in detail. The study focuses on the recent growth of data science in solving various exponential problems, the capabilities of the field in early prediction, measurement of spread, and its outcomes. Finally the study established an optimistic look toward the capabilities of data science in future outbreak of similar viruses. This study does not reflect an exact solution but suggested many underlying mechanisms through which guided study in data science can be performed on COVID-19 and similar outbreaks.
References

1. Khetrapal, S., Bhatia, R.: Impact of COVID-19 pandemic on health system & sustainable development goal 3. Indian J. Med. Res. 151, 395 (2020)
2. Surveillances, V.: The epidemiological characteristics of an outbreak of 2019 novel coronavirus diseases (COVID-19)—China. 2020. China CDC Wkly. 2, 113–122 (2020)
3. World Health Organization. Pandemic influenza preparedness and response. World Health Organization. Accessed 23 Dec 2009 (2010)
4. Sujatha, R., Chatterjee, J.: A machine learning methodology for forecasting of the COVID-19 cases in India 2020. Stoch. Environ. Res. Risk Assess. 34, 959–972 (2020)
5. Yang, Z., Zeng, Z., Wang, K., Wong, S.-S., Liang, W., Zanin, M., et al.: Modified SEIR and AI prediction of the epidemics trend of COVID-19 in China under public health interventions. J Thorac. Dis. 12, 165 (2020)
6. Sivarajah, U., Kamal, M.M., Irani, Z., Weerakkody, V.: Critical analysis of big data challenges and analytical methods. J. Bus. Res. 70, 263–286 (2017)
7. Feldman, B., Martin, E.M., Skotnes, T.: Big data in healthcare hype and hope. Dr. Bonnie. 360, 122–125 (2012)
8. Bellazzi, R.: Big data and biomedical informatics: a challenging opportunity. Yearb. Med. Inform. 23, 8–13 (2014)
9. Raghupathi, W., Raghupathi, V.: Big data analytics in healthcare: promise and potential. Heal. Inf. Sci. Syst. 2, 3 (2014)
10. Baro, E., Degoul, S., Beuscart, R., Chazard, E.: Toward a literature-driven definition of big data in healthcare. Biomed. Res. Int. 2015, 639021 (2015)
11. Manogaran, G., Thota, C., Lopez, D., Vijayakumar, V., Abbas, K.M., Sundarsekar, R.: Big data knowledge system in healthcare. In: Internet things big data technologies for next generation healthcare, pp. 133–157. Springer (2017)
12. Mezghani, E., Exposito, E., Driira, K., Da Silveira, M., Pruski, C.: A semantic big data platform for integrating heterogeneous wearable data in healthcare. J. Med. Syst. 39, 185 (2015)
13. Lin, K., Xia, F., Wang, W., Tian, D., Song, J.: System design for big data application in emotion-aware healthcare. IEEE Access. 4, 6901–6909 (2016)
14. Wang, C.J., Ng, C.Y., Brook, R.H.: Response to COVID-19 in Taiwan: big data analytics, new technology, and proactive testing. J.A.M.A. 323, 1341–1342 (2020)
15. Desai, M.A.: Importance of AI & big data in the fight against coronavirus. Tathapi (UGC CARE J.). 19, 455–466 (2020) ISSN 2320-0693
16. Mayo, C.S., Matuszak, M.M., Schipper, M.J., Jolly, S., Hayman, J.A., Ten Haken, R.K.: Big data in designing clinical trials: opportunities and challenges. Front. Oncol. 7, 187 (2017)
17. Pocock, S.J., Hughes, M.D., Lee, R.J.: Statistical problems in the reporting of clinical trials. N. Engl. J. Med. 317, 426–432 (1987)
18. Yuan, J., Li, M., Lv, G., Lu, Z.K.: Monitoring transmissibility and mortality of COVID-19 in Europe. Int. J. Infect. Dis. 95, 311–315 (2020)
19. World Health Organization. Naming the coronavirus disease (COVID-19) and the virus that causes it. https://www.who.int/emergencies/diseases/novelcoronavirus-2019/technical-guidance/naming-the-coronavirus-disease-(covid-19)-and-the-virus-that-causes-it
20. Novel CPERE: The epidemiological characteristics of an outbreak of 2019 novel coronavirus diseases (COVID-19) in China. Zhonghua Liu Xing Bing Xue Za Zhi. 41, 145 (2020)
24. Rao, A.S.R.S., Vazquez, J.A.: Identification of COVID-19 can be quicker through artificial intelligence framework using a mobile phone–based survey when cities and towns are under quarantine. Infect. Control Hosp. Epidemiol. 41, 826–830 (2020)
25. Covid, C.D.C., Team, R.: Severe outcomes among patients with coronavirus disease 2019 (COVID-19)—United States, February 12–March 16, 2020. MMWR Morb. Mortal. Wkly. Rep. 69, 343–346 (2020)
26. Daurat, A., Millet, I., Roustan, J.-P., Maury, C., Taourel, P., Jaber, S., et al.: Thoracic Trauma Severity score on admission allows to determine the risk of delayed ARDS in trauma patients with pulmonary contusion. Injury. 47, 147–153 (2016)
27. Latif, S., Usman, M., Manzoor, S., Iqbal, W., Qadir, J., Tyson, G., Castro, I., Razi, A., Boulos, M.N.K., Weller, A., Crowcroft, J.: Leveraging data science to combat COVID-19: a comprehensive review 2020. https://www.researchgate.net/publication/340687152_Preprint_Leveraging_Data_Science_To_Combat_COVID-19_A_Comprehensive_Review
28. Livingston, E., Bucher, K.: Coronavirus disease 2019 (COVID-19) in Italy. J.A.M.A. 323, 1335 (2020)
29. Rahimian, F., Salimi-Khorshidi, G., Payberah, A.H., Tran, J., Ayala Solares, R., Raimondi, F., et al.: Predicting the risk of emergency admission with machine learning: development and validation using linked electronic health records. PLoS Med. 15, e1002695 (2018)
30. Sohrabi, C., Alsafi, Z., O’Neill, N., Khan, M., Kerwan, A., Al-Jabir, A., et al.: World Health Organization declares global emergency: a review of the 2019 novel coronavirus (COVID-19). Int. J. Surg. 76, 71–76 (2020)
31. Cockburn, I.M., Henderson, R.M.: Absorptive capacity, coauthoring behavior, and the organization of research in drug discovery. J. Ind. Econ. 46, 157–182 (1998)
32. Kubinyi, H.: 3D QSAR in drug design: volume 1: theory methods and applications, vol. 1. Springer (1993)
33. Murphy, R.F.: An active role for machine learning in drug development. Nat. Chem. Biol. 7, 327–330 (2011)
34. Li, L., Yang, Z., Dang, Z., Meng, C., Huang, J., Meng, H., et al.: Propagation analysis and prediction of the COVID-19. Infect. Dis. Model. 5, 282–292 (2020)
35. Abidi, S.S.R., Goh, A.: Applying knowledge discovery to predict infectious disease epidemics. In: Pacific Rim Int. Conf. Artif. Intell., pp. 170–181. Springer (1998)
36. Wynants, L., Van Calster, B., Bonten, M.M.J., Collins, G.S., Debray, T.P.A., De Vos, M., et al.: Prediction models for diagnosis and prognosis of covid-19 infection: systematic review and critical appraisal. B.M.J. 369 (2020)
37. Kraemer, M.U., Yang, C.H., Gutierrez, B., Wu, C.H., Klein, B., Pigott, D.M., Du Plessis, L., Faria, N.R., Li, R., Hanage, W.P.B.J.: The effect of human mobility and control measures on the COVID-19 epidemic in China. Science. 368, 493–497 (2020)