Summer Training Report

By

Pallak Khanna (0950)

MBA (Hospital and Health Management)

2019-21

IIHMR UNIVERSITY

ACKNOWLEDGEMENT

In the times of uncertainty due to the pandemic, COVID-19 it became almost impossible to undergo an extensive Summer Training Internship. I would like to express my immense gratitude to our University, Indian Institute of Health Management and Research, Jaipur Rajasthan, for giving me this opportunity of research work and online course as an alternative to Summer Training Internship. Our university left no stone unturned in helping the students to cope up with the curriculum in these most difficult times and unseen scenarios.

First and Foremost, I would like to thank Dr. P. R. Sodani, Dean, Institute of Health Management Research, IIHMR Jaipur, for his vision, vitality and guidance at every step to bring this Training Report to fruition.

My heartfelt gratitude goes to my Faculty Mentor, Dr. Monika Chaudhary, for her constant support, encouragement and guidance who helped me with valuable pieces of information and suggestions.

I would also like to thank our class representative for enforcing the deadlines and making sure the students conclude the report on time.

It would be difficult to list all the people who gave their ideas and critical insights in the making of this Report. My gratitude goes to all who helped me by being a constant source of enthusiasm, practical wisdom, support and motivation.

TABLE OF CONTENTS

IMPLEMENTATION OF BIG DATA IN PUBLIC HEALTH, TELEMEDICINE AND HEALTH INFORMATICS IN A POST COVID-19 WORLD

1. ABSTRACT

Background: The implementation of Big Data analytics in health sector, particularly in Public health, Telemedicine and Health informatics has tremendous potential for improving the cost effectiveness, quality of care and efficiency of health systems specially in a Covid-19 world.

Purpose: The objective of this systematic review of literature is to determine the role of Big Data tools and technologies in Healthcare along with its applications in fighting COVID-19. It also pinpoints a few challenges and strategies to overcome them.

Data sources: The articles were reviewed after conducting an extensive systematic search on various major scientific database platforms: ScienceDirect, Emerald, PubMed, IEEE Xplore, Taylor & Francis. The articles that were shortlisted for review were published in English language between 2015 to 2020.

Study selection: Comprehensive articles and replicable studies involving Big Data analytics in public health, telemedicine and health informatics and Big Data in healthcare in a Covid-19 world were selected.

Data extraction: The secondary data extracted includes material on the Role of Big Data analytics in a post COVID-19 world; sources and techniques of Big Data in healthcare; applications of big data in public health, telemedicine and health informatics; implications and relevant strategies to overcome them.

Results: A total of 50 articles were covered in the review according to the inclusion criteria. When the included articles were studied further it was found that: (1) Big data Analytics is proving to be of constant assistance in the fight against COVID-19. Sources of Big Data in healthcare includes private health organizations as well as government hospitals, laboratories, biotechnology and pharmaceutical companies, medical journals, social media and personal monitoring devices. Machine learning coupled with AI in healthcare is achieving new heights in terms of health data management. (2) Immense potential has been seen in the practical use of Big Data analytics in the field of Public health, telemedicine and health informatics especially while tackling the pandemic. (3) Big Data Analytics in healthcare is ensuring highest standards of quality care, patient satisfaction and value based healthcare, despite the major challenges it faces.

Conclusion: This systematic review study unveils the utilization of Big data Analytics in the real-time handling of COVID-19 pandemic. It includes the sources, tools, technologies, applications and challenges of applying Big Data Analytics in Healthcare sector. The usability studies that are replicable are the ones have been taken into consideration. Most of the studies included in the review were from developed countries which necessitate the promotion of research and development of Big Data analytics in healthcare framework in the developing countries for handling COVID-19.

2. INTRODUCTION

The volume of data gathered and analysed by hospitals and public health organizations globally was already booming, but the worldwide pandemic has brought the world of data science under a limelight in a way that has never been possible before.

With patterns changing not day by day, but hourly during the pandemic, healthcare experts battle to provide better diagnosis and quality of care, cost effectiveness, maintain health records, post treatment monitoring and improve each accessible asset. Simultaneously, decision makers must acclimatize new exploration discoveries, alter arrangements, and adopt progressive ways on the grounds of acting rapidly in the times of medical crisis. During this global healthcare emergency, use of data can assist healthcare organizations, hospital, pharmaceutical companies and insurance companies to come up with the necessary solutions to benefit the common man.

According to Mckinsey, Big data refers to the datasets when their size is far beyond the capacity of average database programming software to capture store analyze and manage

Big data is a heterogenous mix comprising of organized, semi structured as well as unstructured data gathered by various organizations globally that can be mined for valuable information which can be utilized in AI ventures, predictive modelling and various other advanced analytics applications. Organizations utilize the enormous information aggregated in their frameworks to improve operations, give better client support, prepare personalised marketing campaigns dependent on explicit client inclinations, which eventually leads to the increment of profitability. Organizations that leverage big data hold a competitive advantage over those that don't since they are equipped for strong and efficient decision making if they utilize the data effectively. [1],[2]

The characteristics of Big data are defined as 5Vs:

- Volume: The quantity of Big data
- Velocity: The speed of data generated, and data collected
- Variety: The different types of data collected
- Veracity: The sources that affect accuracy of data collected
- Value: The benefits derived from data

An enormous part of the relevant healthcare data exists in unstructured or semi-organized structure. The unpredictable, dynamic and heterogeneous attributes of the data makes the extraction of valuable information extremely difficult when traditional data analytical tools and techniques are applied. Thus, there is a limited human capacity to process this information without compelling decision support. This demands the requirement for integration of Big Data analytics into healthcare sector.[2],[8]

COVID-19's rapid spread and effect have caused individuals around the globe to feel vulnerable, frightened, and baffled. In a post COVID-19 world clinical experts, scientists and healthcare providers, are relying heavily on data to have accessibility to updated, real-time information sources for improved and worthy insights. Thus, Big Data is an asset that is helping estimate and comprehend the range and effect of coronavirus. It is being utilized by healthcare professionals, researchers, epidemiologists, and policymakers to sum up and incorporate big data on a local and worldwide scale. While expressions of social distancing, lockdowns and flattening the curve are turning out to be regular sayings in a normal man's life, Big Data is proving to be resourceful in estimating the metrics about population movements across districts, checking public compliance in health protocols, data of the individuals showing symptoms, recovery rates, mortality rates, etc., all of which helps predicting how the curve grows or flattens during a specified period of time. Big data analytics has proved to be a great advantage and a blessing for handling the pandemic.[10],[13]

The timing is immaculate as the healthcare infrastructure during previous epidemics did not have access to Big data analysing tools. The more we know and find about this virus, the more are opportunities to defeat it. This implies more data and a need to protect it as well. According to experts, plethora of information (particularly location data) makes it vulnerable against cyber-attacks. Since we cannot deny the growing significance of Big Data, we need to work out on approaches to protect it as well. The blend of Big Data,

Internet of Things, AI, machine learning, cloud computing and other data-driven advances in the crisis can significantly improve our comprehension of the disease, expand accessibility to healthcare services, and help us share data in real-time globally.[10],[11],[29]

Time is an extravagance that global leadership does not have as COVID-19 has overturned the usual state of affairs for governments and organizations, which should now struggle to protect both lives and jobs. Governments and medical professionals, are working relentlessly to minimize casualties and contain an infection that as of now has spread among millions and claimed plenty of lives. Healthcare management around the world face a multifaceted test: battling the onslaught of the pandemic on the vanguard while additionally handling identical problems as any other industry, for example, worker health and monetary distress.[27],[24],[29]

In the current unfamiliar scenario, where the challenges keep on evolving, it's not astonishing that Big Data analytics, generally recognized for its critical analysis and predictive ability, has become a basic navigational tool for organizations and governments alike. What is astonishing is how swiftly different organizations and governments, even those with limited Big Data experience, have evolved and incorporated different analytic tools to enhance their response and day to day operations. Big Data capacities and infrastructure that would have taken years to develop, were built up from scratch in a matter of few weeks.[23],[27],[30]

The promising performance of Big Data technology in healthcare services has established deep interest among students, researchers and industry investigators. However, there are only a few literature reviews available regarding big data in healthcare industry in dealing with the pandemic COVID-19, which seems to be largely fragmented. Therefore, this study aims to offer a comprehensive understanding on the scope of big data implementation in health care systems including it's applications and implications when adopted in a world post COVID-19. It answers the research question: “How Big Data analytics is implemented in healthcare environment to enhance care delivery with applications in public health, telemedicine and health informatics in a COVID-19 world?” This review also illustrates the implications of leveraging Big Data analytics in healthcare and the relevant strategies to overcome them.

3. METHODOLOGY

This Systematic review was done by collecting and analysing relevant literature from various sources, which focussed on the objectives:

- To determine how Big Data is helping in the fight against COVID-19 by incorporating contemporary analytical technologies
- To illustrate the applications of Big Data in handling COVID-19 in the disciplines of:

    1. Public Health
    2. Telemedicine
    3. Health informatics

- To identify the challenges posed in Big Data implementation in Healthcare and strategies to overcome them.

3.1. Information sources

The articles were reviewed by conducting an extensive systematic search on various major scientific database platforms such as: ScienceDirect, IEEE Xplore, Emerald, PubMed and Taylor & Francis. The references mentioned in the selected publications were also studied for a detailed review.

3.2. Selection criteria

The following inclusion criteria was considered while selecting the literature for this systematic review:

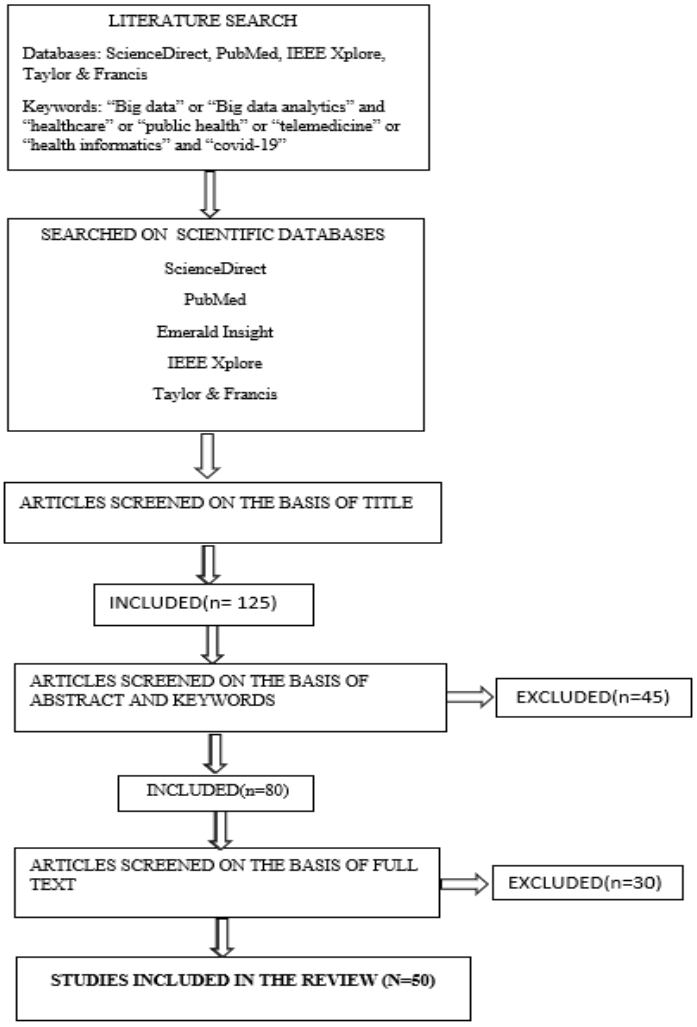
- IC 1: Articles that deal with Big Data analytics in Health domain.
- IC 2: Articles that deal with Big Data applications in handling COVID-19
- IC3: Articles where the primary emphasis is on contemporary data analytics techniques and relevant technology.
- IC 4: Articles published between 2015 and 2020.
- IC 5: Articles published in English Language.

3.3. Study selection

The literature search and selection of research material was carried out in the following steps (Fig. 1):

1. A systematic search of research publications on scientific databases with keywords such as “Big data” or “Big data analytics” and “healthcare” or “public health” or “telemedicine” or “health informatics” and “covid-19”
2. Relevant materials were chosen in accordance with the selection criteria by auditing and identifying important articles using titles and keywords.
3. Research materials that were shortlisted in the previous phase were re-examined for the review.
4. Cross-referenced material was searched and documented for a thorough analysis.

Figure.1. Flow Chart Of Methodology



4. RESULTS

The selection of articles to be reviewed was done by conducting a search on the major Scientific databases. Search results were yielded in large numbers on the platforms, PubMed, ScienceDirect, IEEE Xplore and Taylor & Francis. The studies selected were based on the inclusion criteria. The full text of the selected Articles were read. The literature included in this systematic study comprised mainly of articles which were comprehensive, replicable and methodological. The data from these articles was extracted and analysed. The interpretations of this data is organized into three different groups: The Role of Big Data in the fight against COVID-19 and the technologies used, Applications of Big Data in Public health, telemedicine and health informatics in a post COVID-19 world and lastly, the challenges of big data implementation in healthcare and the relevant strategies to overcome them. The following section summarizes the interpretation and analysis of the findings in each of the objective:

4.1.HOW BIG DATA IS HELPING IN THE FIGHT AGAINST COVID-19

SARS-CoV2 a novel coronavirus, which is responsible for the pandemic COVID-19 as declared by the World Health Organization, has led to the generation of an unprecedented amount of data. Owing to the latest technologies in molecular and computational field of ICTs and Artificial Intelligence, this enormity of information is being managed effectively. For the visualization of real time data, computational techniques are being utilised. One such application was designed at John Hopkins University, USA.

Big data that is collected and analysed from social media networks around the world, gives us an early epidemiological report of the outbreak. This can be seen in the observational study of population level performed by Sun and colleagues [44],[45] for the monitoring of health-related websites, news reports and social media networks in January 2020 in mainland China. It also illustrated how non-conventional datasets can help broaden the understanding of researchers in terms of health resources utilization, health literacy and healthcare seeking behaviours. Effective healthcare measures can be designed and implemented by non-classical data streams. Similarly, in a different study Qin and co-workers [45] utilized Big data in the prediction of new cases of coronavirus, suspected as well as confirmed, by using social media search indexes (SMSI) for keywords including the clinical symptoms of COVID-19. Additionally, by implementing techniques of subset selection method, identified and suspected cases of COVID-19 could

be detected 6-10 days in advance. Whereas Yang and colleagues conducted a research on population migration data to reproduce a Susceptible-Exposed-Infectious-Removed" (SEIR) model, which when combined with Artificial Intelligence based on SARS data to predict the curve of the COVID-19 pandemic in China. They deduced a second peak of infections in the Hubei province of China by mid-March until late April. [3],[7],[8],[9]

Regarding the pandemic, Hua and Shaw [52], insisted that "a unique combination of strong governance, strict regulation, strong community vigilance and citizen participation, and wise use of big data and digital technologies" will lead to the curbing of virus spread in China. Furthermore, the research shows the implementation of various sophisticated technologies which can simultaneously ensure high standards and preserve data privacy. Big data is constantly aiding in the fight against COVID-19 from the initial stages. BluDot, a Toronto-based start-up, had detected unusual pneumonia-like cases in the city of Wuhan, in December 2019. They obtained this information by using Big data Algorithm that acquired data from various sources. The data analysed by the algorithms was extracted from health records, new reports, airline ticketing data and government notices to predict the upsurge of the disease. Airline ticketing data provided a strong base for the prediction of geographical disease spread from Wuhan to other Asian cities. Mobile phone data was also used to track the spread of the virus by analysing location statistics. Big data tools can also inspect the data of senior citizen who are at a risk of contracting the virus. Google's DeepMind artificial intelligence makes use of big data analytics to understand the characteristics of Coronavirus.[12],[13]

In the future, big data will play a major role in preventing pandemics by analysing the global data about the emergence of new viruses or other pathogens, along with disease modelling and visualization of data.[49],[50],[41],[51],[54]

BIG DATA TOOLS AND TECHNOLOGIES IN HEALTHCARE

Big Data in healthcare's enormous size and heterogenous nature renders it relatively less relevant to use the conventional technologies. The commonly used platforms for operating the framework of the software that aids in big data analysis are high power computing clusters which can be accessed via grid computing infrastructures. One such System is Cloud computing which provides digitalized storage technologies and reliable services. Such platforms tend to receive data from sensors universally, further analyses and interprets the data as well as provides the user easy understanding of web-based

visualization. The amalgamation of Internet of things, machine learning and AI approaches for big data can provide the services of mobile edge computing cloudlets and fog computing. Advanced algorithms can be incorporated using ML and AI for big data analysis on computing clusters. The most common programming languages for big data analytics are Python and R, along many others and the most common working platforms for big data analytics are Hadoop and Apache Spark (open-source distributed applications) [1],[2],[3],[4],[5],[6]

Hadoop

MapReduce Framework is implemented by Hadoop for the processing and generation of large datasets to parallelize the computation on multiple nodes. It reduces failures, establishes inter-machine communications across large clusters of machines. With the help of Hadoop and MapReduce, the data is used efficiently which is otherwise impossible to handle.

In various literatures [1],[2], applications of MapReduce and Hadoop were supported by other researchers which showed MapReduce had the ability for improving the performance of common signal pharmacovigilance at linear speedup rates. It can also refine protein structure alignments, improve performance of neural signal processing and increase the speed of medical imaging.

Apache Spark

Another open source application which functions as a consolidated engine for supporting SQL queries(Spark SQL),streaming data(Spark streaming),Machine learning(MLlib) and graph processing(GraphX). It offers productivity of the developer, scalability, fault-tolerance capacity and complex computational ability.

Some sources of primary data collected for healthcare is summarised in Table. 1. The techniques for Big Data Analytics are shown in Table. 2

Table.1. Sources of Health data [27],[31],[32]

S. No.	Sources	Description
1.	-omics/Biological	Whole exome profiling, metabolomics used for etiologic research and screening
2.	Clinical	• Electronic Health Records-detailed information about the patient including prescriptions, medications and medical history • Diagnostics – diagnostic results including imaging and laboratory reports • Biomarkers Ancillary-molecular data and administrative data
3.	Geospatial	Neighbourhood characteristics used in etiologic research and surveillance
4.	Claims	• Medical claims- Medical reimbursement data • Prescription claims-Prescription reimbursement data
5.	Clinical Research	Clinical trials and designing parameters like compound, size and end points.
6.	Personal monitoring	Patient related data generated by daily GPS records and mobile health applications.
7.	Effluent data	Google search results as well as other social media search results used for surveillance and screening of hidden social networks.

Table.2. Big Data Analytical Techniques in Healthcare [1],[2],[32],[33]

Big Data Analytical Techniques	Healthcare Applications
Cluster Analysis	Determination of cluster for identifying high-risk groups
Data Mining	Determination of epidemics and pandemics, Bio signal monitoring and exploratory data analysis in HC.
Graph Analytics	Hospital performance analytics to ensure quality
Machine learning	Predictive analysis, assessment of hospital performance, determination of epidemics/pandemics.
Natural Processing Language	Epidemiology: morbidity, mortality, risk assessment; Efficiency of care and cost, improve treatments, extracting information from clinical records
Neural Network	Diagnostics and predicting natural history of disease
Pattern Recognition	Improvement of public health surveillance
Spatial Analysis	Visual, spatial and advanced analytics to study population health

4.2. APPLICATIONS OF BIG DATA IN HANDLING COVID-19

I. BIG DATA IN PUBLIC HEALTH

Over the years Big Data has been extensively utilized to drive innovations, creating insights and allowing new interventions across all economic sectors. The decrease in costs of technology is contributing to this wide emergence of Big data in Public Health. The use of Big Data in Public health has been shown to improve upon various disciplines such as disease surveillance, predicting risk, targeting interventions, understanding overall population health as well as precision public health. Big data can play a key role in both research and intervention activities by the development of the databases on the prevalence and incidence of diseases. Other examples of big data applications in public health along with epidemiology includes maternal and child health, environmental health science, nutrition, clinical prediction and diagnostics. Big data in public health, thus can promise to offer value-based healthcare.

Due to the COVID-19 outbreak and a subsequent rise in cases at an alarming rate worldwide there is a gigantic surge of data from all around the world including the number of cases, travel history and records of co-morbidities. This calls for immediate attention to capture and analyse data for current and future interventions. Thus, scientists and healthcare professionals are making use of sophisticated technologies such as, big data analytics, machine learning and natural language processing for tracking Coronavirus. Big Data can be used as a powerful tool in storing and analysing these large datasets and identifying trends, predicting patterns, associations and differences that can help in the detection and recovery of COVID-19 patients [3]. With the capability of detailed data capturing, big data can be beneficial in minimising the risk of spreading Coronavirus, case identification and allocating resources for better treatment. Data such as patient location, proximity, travel history, symptoms can be digitalised and processed for generating much needed insights for interventions both at community and global level.[6],[7] The following table (table.3.) provides applications of Big Data in Public health for handling COVID-19:[6],[7],[13],[23],[29],[51]

Table 3. Applications of Big Data in Public health for COVID-19 handling

S.No.	Applications	Description
1.	Identification of infected cases by Testing	Sample collection of suspected patients or in case of random mass testing, information of all positive, negative or invalid samples are stored. Storage of complete medical history of all patients.
2.	Symptoms	Digital record of the symptoms, ranging from mild to severe to critical. Helps to identify patients who need emergency care.
3.	Travel History	Travel records of the suspected and identified cases are stored for the prevention of spread by determining the mandatory quarantine/isolation periods, also helps in contact tracing.
4.	Detection of Coronavirus at an early stage	Rapid detection leads to the adoption of preventive measures.
5.	Information during lockdowns	Tracking and monitoring the data of the patients during lockdowns.
6.	Migration	People entering or leaving the affected city or hotspot. With this data suspected cases can be identified.
7.	Development of new treatments	Rapid growth in research and development of medicines, medical devices and equipment needed for current and future healthcare needs.
8.	Availability of resources	Monitoring the availability of hospital resources and bed capacity.

As shown in the table, big data analytics acts as an efficient contraption for tracking, controlling, research and prevention of COVID-19 as a pandemic.

In China COVID-19 was suppressed with the help of implementing data collection with Artificial Intelligence which led to a decrease in the rate of spread. A research team from the University of Hong Kong HKU came up with a method to track the spread of COVID-19 using population flow data in real-time using big data analytics. The team utilized information from real-time migrations centred around aggregate population flow. Any cell-phone user that has spent a minimum of 2 hours in Wuhan during the period of

the study was taken into account in the dataset. When the users turned their phones on, their geographical location was recorded. The population flow statistics were then combined with the geographical location and the total number of COVID-19 cases up until 19 February 2020 in China. By doing this, researchers proved that the respective quantity of migration from Wuhan, the epicentre, precisely anticipated the comparative density and spread of COVID-19 cases around China.[8],[10],[11],[12]

Big Data will help in fast-tracking vaccine development based on more absolute knowledge on the virus and better insight on how it spreads. For developing comprehensive and valuable antibody testing, and to enhance vaccine development, organizations have started leveraging advanced big data analytics tools. As of now, there is a lot of uncertainty about the level of immunity in people after recovering from COVID-19. Big Data analytics and technologies can play a huge role in helping in understanding the degree of immunity and how it shapes the future of the pandemic. Besides enhancing the drug development, data analytics tools is playing a major role in alleviating the spread of the virus. As most countries are starting to reopen their economies, big data along with other IT and analytical tools will be essential in averting a second wave of coronavirus infections that could overburden hospitals and health systems globally.

Several organizations are utilizing these technologies in accelerating COVID-19 drug discovery and figuring out how our immune system fights the virus. In April 2020, the pharmaceutical giants GlaxoSmithKline (GSK) and Vir Biotechnology partnered up to make headway towards the development of a viable COVID-19 treatment by intersecting Artificial intelligence and CRISPR. Several organizations like Amazon web services and Google cloud have launched projects which offered free access to researchers to open datasets and analytics tools to help them develop rapid solutions for COVID-19. Organizations like Definitive Healthcare has partnered up with Esri in USA and launched an interactive data platform to analyse and monitor hospital bed capacity, availability of resources and geographical areas of potential risk during the pandemic. Paraxel, one of the largest clinical research organization in the world, is attempting to create large scale COVID-19 RWE(Real World Evidence) studies that can extract data from patients currently undergoing treatment and is actively working to form a coalition of drug development experts and industry associations for the same. Also, Google, the behemoth

in the world of search engines naturally has access to a massive amount of data which it had already given access to the Centre for Disease Control(CDC) in 2015 to predict the spread of preventable STDs. The company had announced in 2019 that it has been using its AI systems to detect lung cancer in its early stages by reviewing CT scans from National Cancer Institute and North-western University.[12],[13][40],[41],[45],[46],[47],[48],[50]

With the amount of data being generated due to the outbreak of the pandemic it seems like a tailor-made situation has been created for advanced analytics technologies to become a part of our daily lives. To achieve a breakthrough in this situation, we need to bring together professionals who have the skillset of advanced computation and Artificial intelligence, and those who understand population models and health systems. Handling of the pandemic is an interdisciplinary problem, and to defeat this deadly pathogen, we need the data and we need the merger of healthcare and IT.

II. BIG DATA IN TELEMEDICINE

Telemedicine can be referred as an enabler of healthcare accessibility and affordability because it aids in the delivery of health care services by health care professionals all around the world by using communication and information technologies for the exchange of valid diagnosis, curative and preventive healthcare, research and development, and for imparting health education, all in the interests of enhancing community health. In other words, telemedicine is the remote diagnosis and treatment of patients with the help of electronic telecommunication technologies which allow patient and doctor contact, care, advice, reminder, education, intervention and monitoring.

There are a lot of potential benefits of telemedicine. It speeds up the interventions by providing faster accessibility to services that may not be otherwise available. In a country like India, providing person to person healthcare is rather challenging due to the insufficient resources and large geographical distances. One major advantage served by telemedicine is the reach of medical services in rural areas which also saves up the cost and effort of the patients.[14],[16],[17]

While the global economy is in the throes of an unprecedented recession in recent memory, the COVID-19 pandemic has created a massive disruption around the world putting the healthcare industry and all the major industries under a huge stress. Entire governments are reorganizing and mobilizing on a scale that hasn't been seen since the World War 2 to tackle the pandemic. In such circumstances, when the healthcare infrastructure around the globe is fatigued and wearing down, it has become imperative to find ways to reduce the stress on our healthcare professionals while at the same time increasing their efficiency. Telemedicine with lesser costs, better penetration and a farther reach has become the need of the hour.

Sparked by the coronavirus pandemic, telemedicine has seen an enormous upsurge in its market across the globe. As more and more healthcare providers take up telemedicine and telenursing to fight COVID-19, this segment powered by Big Data analytics is capturing a previously untapped market and completely transforming how healthcare happens. As per a recent report by Global Market Insights Inc, the present market value of telemedicine is around 31.8 billion dollars and is expected to reach $41.2 billion by 2021 and $130.5 billion by 2025.

How is Big Data in telemedicine transforming healthcare? [16],[19],[21],[22],[25],[26]

1) Using the cloud for data collection and analysis

The last decade has seen the transition of medical records from paper based to electronic data stored on the cloud which has led to an increase in information available to big data analysts to uncover patterns, correlation and trends. The boom in telemedicine practice is creating a huge amount of such data as medical information needs to be shared virtually , especially when the patient's treatment is being done remotely.

2) Remote monitoring, medication and post-treatment follow up

The availability of smart health devices employing Internet of Things(IoT) gives healthcare providers access to patient's vitals, also generating a humongous amount of medical data. The catalogue of medical offerings on telemedicine platforms help treat patients that are restrained by cost constraints, geography and physical limitation. Remote monitoring ensures that the public can utilize healthcare services from the comfort of their homes when required. Effective utilisation of real time patient data and its processing using big data tools helps to compute required medication dosages and help

develop treatment option. When the treatment is completed, monitoring of patient's health is done remotely to avoid follow-up visits with doctors. This is especially useful in the case of elderly or vulnerable patients or those that have been treated for a highly infectious disease and may still be carrying the pathogen.

3) Predictive analysis and predicting trends

Predictive analysis using big data techniques help in identifying individuals who are at risk of developing certain ailments and health conditions. Big Data lets health analysts create individual risk scores as per the information collected and analysed. This helps in identifying illnesses at early stages of progression leading to better treatment outcomes. Predictive analysis is especially beneficial if the patients have complicated chronic diseases. Doctors can predict medical emergencies in advance and avoid deterioration of health of their patients. Big data is instrumental in creating disease spread models. Applying artificial algorithm to big data problems in healthcare helps in predicting infection trends and disease progression. This gives the healthcare providers an insight into how an infection spreads through a population group and helps in identifying regions that are most likely to be affected during an outbreak.

This is especially useful in the current scenario where the healthcare infrastructure is stretched to the brink in the fight against COVID-19. Telemedicine services are then employed to enable medical professionals to provide efficient healthcare to a large population within the identified region.

4) Precision medicine

By applying big data analytics, telemedicine treatment options can be personalized to fit the requirement of individual patients. Individual data can be extracted from wearable devices, fitness and health apps, old medical records and used in developing treatment catering to the specific needs of each patient. This idea is a step ahead from the traditional approach of general medicine where a single type of treatment was prescribed to all individuals suffering from a certain ailment without taking into account their lifestyles, environmental conditions and genetic makeup.

5) Reducing Risk during Pandemic

The Covid-19 pandemic tends to pose unique challenges in providing healthcare. Thus, telemedicine is well suited for such scenario where contactless healthcare is provided.

Telemedicine can be utilized without exposing the medical staff and the patients to infections in the time of outbreaks. The transmission of infectious diseases such as COVID-19 can be prevented by practising telemedicine as it will reduce the risks both to the patients as well as healthcare workers. However this practice is only apt for patients with mild symptoms. In this case, Big data analytics can help in digital storage and evaluation of data and improve decision making by healthcare professionals. By combining the power of telemedicine and big data, the number of unnecessary OPD visits is significantly reduced without any compromise on outcomes.

6) Drone applications for disasters and pandemics

Drones collect and use data for navigation and performing preliminary surveys after a disaster or the area of the outbreak of disease. Drones are used for delivery of medical supplies and medical equipment to areas with limited human reach. Drones based telemedicine can be specifically useful for the urgent delivery of life-saving vaccines or antidotes. Next-gen start-ups are using drones and big data analytics to deploy medical cargo drones in several countries to tackle COVID-19. With the help of delivery drones, Covid-19 test samples from rural areas are taken to urban laboratories and medicines are transferred from urban to the rural areas.

In the current scenario of COVID-19 outbreak, healthcare professionals are regularly getting infected by the disease which leads to decrease in patient-doctor ratio. Telemedicine might be the only alternative of traditional medical practice and promising solution available specially with the adjunct of big data which makes it a powerful technology. However, like any other technology, telemedicine has it’s risks, limitations and drawbacks which can be diminished with proper training and enforcement of standard protocols and guidelines.

III. BIG DATA IN HEALTH INFORMATICS

Health Informatics, also commonly known as healthcare informatics, is a field of information engineering applied to the health domain, essentially for the management of patient care statistics. It is a multidisciplinary field that utilizes Health Information Technology (HIT) to improve health care by delivering high standards of quality and efficiency, to incur low costs, greater availability and new opportunities. According to the United States National Library of Medicine it can be defined as "the interdisciplinary study of the design, development, adoption and application of IT-based innovations in health care services delivery, management and planning'. Thus, health informatics can be classified as a diverse field which cannot be encapsulated into a single definition. As the healthcare services are collectively referred to as "there is no one size that fits all", healthcare professionals are using health informatics with the help of HIT and data management to provide precision healthcare.

The sub-fields of health informatics include: Clinical Informatics, Imaging Informatics, Nursing Informatics, Public Health Informatics, Medical Informatics, Pharmacy Informatics, Bioinformatics and Translational Bioinformatics. Health informatics including all the subfields has been revolutionizing the traditionally practices of health care, from caring of each patient individually to improving the health of entire population. It is a continuously evolving science with the expansion of Electronic Health Records along with Health Data Analytics System. With the establishment of clinical terminology sets and health data exchange standards, health informatics has grown tremendously. As seen in the recent time Health Informatics is one of the world's fastest growing sector, with people who possess the right skillset, training and experience.[27],[28],[33],[34],[13]

Health informatics, as a field of Computer science and Information technology within the realm of healthcare, is on the cusp of a breakthrough while it's entering a new epoch, where the advancements in technology makes it capable of handling Big Data. Big data in Health Informatics shows immense potential for unlimited growth of information. Data Mining techniques and Big data in health informatics aims to focus on the ultimate goal of providing improved Health Care Output (HCO). In other words, it aims to deliver value-based healthcare by improving the quality of care to the patients.

While the number of cases of COVID-19 is on a constant rise, health systems around the world are promptly anticipating and adjusting to the expanding clinical demands. Inherent

to the operational administration of a pandemic in the time of advanced medication, the capabilities of Electronic Health Records(EHR) are being leveraged, which can be helpful for the development of devices to aid the standard management of patients. Technology-based devices can adequately provide support to the healthcare organizations and hospitals during a pandemic by encouraging immediate dissemination of data, tracking transmission of the virus progressively, creating virtual venues for gatherings and everyday tasks, and above all, offering telemedicine visits for patients. However, at the time of Ebola outbreak in 2014, consideration was brought to the utilization of the EHR as a potential public health tool. Unfortunately, regardless of the ongoing Ebola pestilence, proper infrastructure for the pandemic were absent in the US health systems along with their EHR applications. As we are currently confronting a pandemic, numerous establishments are attempting to rapidly develop steady and solid informatics framework so as to get ready for dealing with an exponential increment in Coronavirus infected patients.

The role of biomedical big data, AI and advanced informatics play an important role in managing a disease outbreak. In handling COVID-19, Big data can be applied in the following subfields of health informatics:[33],[34],[29],[38]

- Translational bioinformatics: the data can be useful in the development of targeted drug and vaccine by investigating the mechanism of SARS-CoV-2.
- Sensor informatics: the data from the wearable sensors can be extracted for the real-time monitoring of asymptomatic and home-based patients of COVID-19.
- Imaging Informatics: the data from the imaging reports such as CT or X-ray reports of the COVID-19 patients can be utilized effectively along with RT-PCR to improve accuracy of COVID-19 diagnosis.
- Clinical Informatics: The multimodality data is analysed to help in hospital management by using an effective clinical workflow to ensure high standards of quality care for critically ill COVID-19 patients.
- Behavioural informatics: Human behaviour data such as self-quarantine, community quarantine etc. is analysed to prepare better policy and execution routine.

- Mental health Informatics: Anxiety levels, emotional health and overall mental health data of the patients is analysed to provide relevant solutions for the people to help them cope up with any potential fear or restlessness due to the pandemic.
- Rehabilitation Informatics: Physiological data is analysed to predict the full recovery of all the organs of different patients suffering from different level of infections (mild to critical).
- Public Health Informatics: the field data of the disease is analysed for proper population health management and COVID-19 care resource supply chain management.

The health care information mainly functions on the digital databases which are Electronic Health Records and serves as the digital version of health information systems. EHR includes all the information on diseases, consultations and test results which allow patients and health care providers to easily share the medical and clinical data for ensuring quality care. EHR makes the patient information more accessible across all the departments of the health care facilities. With the help of EHR, everything is centralized including, diagnosis, follow-ups, treatments, possible allergies, diseases and vitals. With the current scenario, courtesy of the pandemic, a steady stream of medical data is being generated. Thus, with the integration of Big Data Analytics in EHR, the pandemic can be managed efficiently. Table.4 shows the role of big data in EHR in handling COVID-19.

Table 4. Role of Big data in EHR in handling COVID-19 [7],[9],[11],[15],[28],[29]

S.NO.	Applications	Description
1.	Screening	Triage of Patient Phone Calls Questionnaire for Screening Patients Records of Physical Screening of symptoms
2.	Patient management	Managing the concerns of the patients regarding care providers. Instructions for home Isolation/Quarantine

3.	Inpatient department, emergency care and ambulatory care	Data of the patients according to severity of infections and type of care needed Number of PPE needed for health care providers. Data of the total specimen collected.
4.	Reporting and Analytics	Tracking of COVID-19 cases in an EHR embedded database Tracking the suspected cases or contact tracing in EHR databases. Data regarding the existing and pending tests, screening or documentation compliance by the patients.
5.	Channels of communication	All the messages in the EHR integrated secure messenger.
6.	Patient Facing Technology	Utilization of the data in EHR to execute video visits for home based COVID-19 patients.

Big data Analytics can be applied effectively to the field of health informatics since it can gather huge amount of heterogenous real-time data, especially in the form of EHRs. The most significant criteria determining the success of any health system in world is Patient satisfaction. Big Data with its advanced set of techniques of evaluation and visualization can help to accomplish highest level of patient satisfaction. Despite of the increasing challenges and issues faced by the patients, big data in health informatics can provide sophisticated and modern administrative services and can greatly impact the future of healthcare sector.

4.3.CHALLENGES OF BIG DATA IMPLEMENTATION IN HEALTHCARE AND THE STRATEGIES TO OVERCOME THEM

Owing to its very nature, big data is extremely complex and massive, which requires the healthcare organisations to implement each step of Big data analysis carefully. Despite its immense potential and added values to the healthcare system, Big Data analytics pose serious challenges in incorporating it's tools and techniques in healthcare. Some typical challenges faced by the organizations and researchers when Big data is implemented in the various fields of healthcare and how they can overcome these challenges to simultaneously achieve data driven healthcare decisions and accomplish the financial goals are as follows:1],[2],[5],[31],[32],[33],[35],[36],[43],[55],[57]

Capture

Challenges: The capturing of data that is clean, accurate, complete and formatted correctly for multimodality is an ongoing conflict for many organizations. In a recent study, at an ophthalmology clinic, when EHR data was analysed it was found that there was a mis-match of the data with under-reporting of the symptoms. The data of the patient which matched with the EHR record was only 23.5 percent. Thus, poor reusability of EHR data, incomplete workflows and inadequate reporting of information can contribute to the improper capturing of quality data.

Strategies: To overcome this challenge of data capturing, healthcare providers can prioritize valuable data type for different projects. They can also integrate the assistance of health information management experts by accommodating data governance integrity. By developing programs for the improvement of clinical documentation, the clinicians would be well trained in handling and reporting data.

Storage

Challenge: The storage of large volume of data constitutes one of the primary challenges.

Strategies: To overcome the storage problem of big data, healthcare organizations use their own IT infrastructure to store the data generated. It improves the accessibility, security and reduces time. However, difficulty in scaling and maintaining on-site server

networks keeps on arising due to the expensive technique. To decrease the cost and increase reliability, most of the healthcare organizations are using cloud based IT infrastructure. Cloud partners must be chosen by the healthcare organizations by making them understand the importance of security issues and specific compliance of health related data. Moreover, cloud storage offers a decrement in cost, easier expansion and data recovery. Many organizations also follow a hybrid approach for data storage and accessibility.

Security and Privacy

Challenges: The world of data science has faced way too many security breaches to count. Health data of an individual is the most intimate personal information that people entrust an organization with. Like other organizations that employ big data analytics, the importance of data security cannot be underestimated in healthcare which has faced its own share of network breaches, phishing and ransomware attacks, and hackings. Once the susceptibilities were observed and studied, a list of rules and technical safeguards, known as HIPAA Security Rules, were put in place for the protected health information (PHI). These rules were meant to help organizations develop safeguards with respect to storage and transmission of data, access control, authorization framework, auditing and verification. Data privacy fixates on the utilization on an individual's sensitive information for purposes that are clearly defined as per policy, whether state's or an organization's own. There are a certain set of guidelines that must be followed when the data is recorded, transferred and used. It must be ensured that data is protected from malicious violations and is not sold for profit.

Strategy: Data security can be enforced by adopting a few basic systems - firewalls, encrypting sensitive data while in storage, end-to-end encryption software, MFA(Multi Factor Authentication) and of course, an updated antivirus software. Data security is essential to safeguard data but it cannot address concerns with respect to privacy. To ensure data privacy, it is imperative that the requisite legal framework is put into place by the concerned authorities and there should be disciplinary action if those guidelines are violated.

Cleaning

Challenge: Right after the acquisition of data, there is an urgent need for cleaning or scrubbing the data to ensure its accuracy, relevancy, purity and consistency. Cleaning of the data is also necessary as the dirty data tends to derail the big data analytics project, specially when encapsulation of heterogenous data is done.

Strategy: Most cleaning or scrubbing processes are done manually, which can take a large amount of time. However some IT vendors offer automated scrubbing tools that use logical processing to improve the accuracy of the data. With the integration of Machine Learning techniques these tools are becoming more precise which impart high quality of data with less costs.

Unified format

Challenge: Patients are responsible for producing a huge amount of medical data, especially with the on-going pandemic, COVID-19, it is not easy to capture data with traditional EHR format as it makes the process tedious and unmanageable. There is a rising need to codify all the medically relevant information including medical claims, bills and clinical analytics.

Strategy: To code all the relevant clinical and medical data, Medical coding systems come into play. These include CPT and ICD code sets. However, these code sets have their own drawbacks.

Image pre-processing

Challenges: Misinterpretations and mis-matched data can be due to various physical factors such as multiple types of artifacts and noises. It usually happens in the medical reports including medical imaging. Improper handling of medical image reports can lead to delineation of anatomical structures and result in wrong diagnosis.

Strategies: By reducing the noise, adjusting contrast of acquired test images, enhancing image quality and clearing artifacts, the errors in handling medical imaging data of the patients can be diminished.

Meta-data

Challenge: Healthcare data must have a long shelf-life for keeping the data of the patient accessible for many years to come. Data needs to be re-reused and re-investigated for research projects which makes stewardship and curation an important concern. Data is also reused for evaluating the quality and the performance of benchmarking. For a successful data governance plan, meta data is a key component. The meta-data constitutes of information about the data such as, time of creation, purpose of the data and person associated with the data and the previous usage of the data by data analysts or researchers.

Strategy: A meta-data steward should be assigned by healthcare organizations to handle the development and the curation of insightful meta-data. A data steward can ensure a properly documented task including, standard formats and all definitions from creation to deletion.

Querying

Challenge: Proper interoperability between the datasets in healthcare organizations is lacking which leads to the generation of incomplete report of medical history of any patient. The entire repository of data is not accessible at once by the query tools. Different components of the datasets are stored in different capsules or different formats, which fails to provide the entire data of the patient altogether.

Strategies: Medical coding systems like ICD-10, SMOMED-CT, or LOINC and Structured Query language (SQL) can solve this problem by querying large and relational datasets.

Visualization

Challenges: Visualization of data is the graphical representation of the data which makes the report look precise and gives an easy understanding of the data for the viewer to absorb maximum amount of knowledge and reduce any confusion. Complex flowcharts, low quality graphics and congested text can cause frustration and misinterpretation of the data

Strategy: For clear depiction of data, different types of charts and graphs are utilized according to the type of data. The examples include heat maps, bar charts, pie charts, histograms, box plots and scatterplots which convey the data in their own specific way.

Data sharing

Challenges: Patients may not receive their care at a single healthcare organization throughout their lives. Thus the nature of data shared from one organization to another may not be interoperable. This is due to the technical barriers between the organizations which leads to a less informed decision making regarding the treatment of patients.

Strategies: Fast Healthcare Interoperability Resource (FHIR) and public APIs, CommonWell and CareQuality can serve as solutions by making the data easily accessible and interoperable. This will also ensure data security.

Ethical and Legal

Challenges: Owing to the personal nature of the data, ethical and legal issues arise in healthcare data. Ethical and legal issues include privacy risks, transparency, autonomy and trust while utilising big data. Other concerns also include the commercialization of data and the usage of medical data against the permission of the people.

Strategies: Any organization using Big data in healthcare should ensure the establishment of affirmative policies to protect the personal data and confidentiality of the patients.

5. CONCLUSION AND FUTURE PROSPECTS

Through this systematic review, it can be presented that Big Data is playing an important role in the fight against COVID-19. It has emerged as a new frontier for the enhancement of Healthcare delivery around the world. A shift from experience based medicine to evidence based medicine can be observed by the implementation of Big Data technologies in healthcare. It is only possible due to the pattern recognition and predictive nature of Big Data. In addition this study includes, Big data definitions, characteristics, tool and technologies, various sources and applications of Big Data in the fight against the pandemic. A clear, but slow transition of the health care sector can be seen from a wider volume based healthcare to precision healthcare. It is imperative to say that Big data is bringing about this transition but it's full potential is only limited to developed countries. It is essential for the healthcare and IT professionals to collaborate and understand the working of Big Data Analytics to bring out a wide technological evolution in health sector. In the coming years, big data analytics can focus towards developing a predictive system in healthcare. This would impart futuristic prediction of outcomes in population health as well as precision health based on the existing data. Big Data can help in the prediction of disease outbreaks or future epidemic or pandemic in advance and assist in the development of targeted drugs or vaccines and initiate early interventions. The implementation of big data in handling of COVID-19 can pave the way for preventing future outbreaks by making data driven decisions for the future.

Despite the dynamic nature of Big Data, it is essential to proceed with caution while utilizing the benefits of Big Data Analytics. It is important to keep in mind that the existing big data models for COVID-19 may be incomplete or there may be the existence of statistical and methodological limitations. Truth be told, Predictive analysis is just predictive and not fortune telling. This means that despite committing our full efforts in Big data analytics, we may still not have the full picture while dealing with COVID-19.

6. REFERENCES

1. Mehta, Nishita & Pandit, Anil. (2018). Concurrence of Big Data Analytics and Healthcare: A Systematic Review. International Journal of Medical Informatics. 114. 10.1016/j.ijmedinf.2018.03.013.
2. Dash, Sabyasachi & Shakyawar, Sushil & Sharma, Mohit & Kaushik, Sandeep. (2019). Big data in healthcare: management, analysis and future prospects. Journal of Big Data. 6. 10.1186/s40537-019-0217-0.
3. F.J. Martin-Sanchez, V. Aguiar-Pulido, G.H. Lopez-Campos, N. Peek, L. Sacchi, Secondary use and analysis of big data collected for patient care, IMIA Yearbook 26 (2017) 1–10, http://dx.doi.org/10.15265/IY-2017-008.
4. Z. Goli-Malekabadi, M. Sargolzaei-Javan, M.K. Akbari, An effective model for store and retrieve big health data in cloud computing. Methods Programs Biomed. 132 (2016) 75–82, http://dx.doi.org/10.1016/j.cmpb.2016.04.016.
5. B. Cyganek, M. Graña, B. Krawczyk, A. Kasprzak, P. Porwik, K. Walkowiak, M. Woźniak, A survey of big data issues in electronic health record analysis, Appl. Artif. Intell. 30 (2016) 497–520 http://dx.doi.org/10.1080/08839514.2016.1193714.
6. Wang CJ, Ng CY, Brook RH. Response to COVID-19 in Taiwan: big data analytics, new technology, and proactive testing. JAMA. 2020 doi: 10.1001/jama.2020.3151.
7. Anshari M, Almunawar MN, Lim SA, Al-Mudimigh A. Customer relationship management and big data enabled: personalisation & customisation of services. Applied Computing and Informatics. 2019;15(2):94–101. doi: 10.1016/j.aci.2018.05.004.
8. Bai Y, Yao L, Wei T, Tian F, Jin DY, Chen L, Wang M. Presumed asymptomatic carrier transmission of COVID-19. JAMA. 2020 doi: 10.1001/jama.2020.2565.
9. Xia W, Sanyi T, Yong C, Xiaomei F, Yanni X, Zongben X. When will be the resumption of work in Wuhan and its surrounding areas during COVID-19 epidemic? A data-driven network modelling analysis. SCIENTIA SINICA Mathematica. 2020 doi: 10.1360/SSM-2020-0037.

10. Haleem A, Javaid M, Vaishya R. Effects of COVID 19 pandemic in daily life. Current Medicine Research and Practice. 2020 doi: 10.1016/j.cmrp.2020.03.011.
11. Amini, S., Gerostathopoulos, I., & Prehofer, C. (2017). Big data analytics architecture for real-time traffic control. In: 2017 5th IEEE international conference on models and technologies for intelligent transportation systems (MT-ITS), 26 Jun 2017
12. Wu C, Liu Y, Yang Y, Zhang P, Zhong W, Wang Y, Wang Q, Xu Y, Li M, Li X, Zheng M, Chen L, Li H. Analysis of therapeutic targets for SARS-CoV-2 and discovery of potential drugs by computational methods. Acta Pharmaceutical Sinica B. 2020 doi: 10.1016/j.apsb.2020.02.008.
13. Haleem A, Javaid M, Khan IH, Vaishya R. Significant Applications of Big Data in COVID-19 Pandemic [published online ahead of print, 2020 May 7]. *Indian J Orthop*. 2020;54(4):1-3. doi:10.1007/s43465-020-00129-z
14. Ohannessian R, Duong TA, Odone A. Global Telemedicine Implementation and Integration Within Health Systems to Fight the COVID-19 Pandemic: A Call to Action JMIR Public Health Surveill 2020;6(2):e18810 https://publichealth.jmir.org/2020/2/e18810
15. Greenhalgh T, Wherton J, Shaw S, Morrison C. Video consultations for covid-19. BMJ 2020 Mar 12;368:m998.
16. Hollander JE, Carr BG. Virtually Perfect? Telemedicine for Covid-19. N Engl J Med 2020 Mar 11.
17. Dorsey ER, Topol EJ. Telemedicine 2020 and the next decade. The Lancet 2020 Mar;395(10227):859.
18. Liu S, Yang L, Zhang C, Xiang Y, Liu Z, Hu S, et al. Online mental health services in China during the COVID-19 outbreak. The Lancet Psychiatry 2020 Apr;7(4):e17-e18.
19. Zhou X, Snoswell CL, Harding LE, Bambling M, Edirippulige S, Bai X, et al. The Role of Telehealth in Reducing the Mental Health Burden from COVID-19. Telemed J E Health 2020 Mar 23.
20. Ohannessian R, Mattos S, Paranhos-Baccalà G, Vanhems P. The role of telemedicine in response to the Zika virus outbreak in Brazil. 2016 Presented at: Zika Summit; April 25, 2016; Paris, France.

21. Ohannessian R. Telemedicine: Potential applications in epidemic situations. European Research in Telemedicine/La Recherche Européenne en Télémédecine 2015 Sep;4(3):95-98.
22. Smith A, Thomas E, Snoswell C, Haydon H, Mehrotra A, Clemensen J, et al. Telehealth for global emergencies: Implications for coronavirus disease 2019 (COVID-19). J Telemed Telecare 2020 Mar 20:1357633X2091656.
23. Paterlini M. On the front lines of coronavirus: the Italian response to covid-19. BMJ 2020 Mar 16;368:m1065.
24. Remuzzi A, Remuzzi G. COVID-19 and Italy: what next? The Lancet 2020 Mar 13.
25. World Health Organization Regional Office for Europe. Strengthening the Health Systems Response to COVID-19. Technical guidance #1: Maintaining continuity of essential health care services while mobilizing the health workforce for COVID-19 response. https://euro.sharefile.com/share/view/sbc0659718fd4c8aa
26. Board of Governors. Telemedicine Practice Guidelines Enabling Registered Medical Practitioners to Provide Healthcare Using Telemedicine. Available from: https://www.Mohfw.gov.in/pdf/telemedicine
27. Dash, S., Shakyawar, S.K., Sharma, M. *et al.* Big data in healthcare: management, analysis and future prospects. *J Big Data* **6,** 54 (2019). https://doi.org/10.1186/s40537-019-0217-0
28. Doyle-Lindrud S. The evolution of the electronic health record. Clin J Oncol Nurs. 2015;19(2):153–4.
29. Chaari L, Golubnitschaja O. Covid-19 pandemic by the "real-time" monitoring: the Tunisian case and lessons for global epidemics in the context of 3PM strategies [published online ahead of print, 2020 Apr 25]. *EPMA J*. 2020;11(2):1-6. doi:10.1007/s13167-020-00207-0
30. Panch, T., Szolovits, P., & Atun, R. (2018). Artificial intelligence, machine learning and health systems. *Journal of global health*, *8*(2), 020303. https://doi.org/10.7189/jogh.08.020303
31. McCue, M. E., & McCoy, A. M. (2017). The Scope of Big Data in One Medicine: Unprecedented Opportunities and Challenges. *Frontiers in veterinary science*, *4*, 194. https://doi.org/10.3389/fvets.2017.00194

32. Mooney, S. J., & Pejaver, V. (2018). Big Data in Public Health: Terminology, Machine Learning, and Privacy. *Annual review of public health*, *39*, 95–112. https://doi.org/10.1146/annurev-publhealth-040617-014208
33. Luo, J., Wu, M., Gopukumar, D., & Zhao, Y. (2016). Big Data Application in Biomedical Research and Health Care: A Literature Review. *Biomedical informatics insights*, *8*, 1–10. https://doi.org/10.4137/BII.S31559
34. Kaggal, V. C., Elayavilli, R. K., Mehrabi, S., Pankratz, J. J., Sohn, S., Wang, Y., Li, D., Rastegar, M. M., Murphy, S. P., Ross, J. L., Chaudhry, R., Buntrock, J. D., & Liu, H. (2016). Toward a Learning Health-care System - Knowledge Delivery at the Point of Care Empowered by Big Data and NLP. *Biomedical informatics insights*, *8*(Suppl 1), 13–22. https://doi.org/10.4137/BII.S37977
35. Wang, X., Williams, C., Liu, Z. H., & Croghan, J. (2019). Big data management challenges in health research-a literature review. *Briefings in bioinformatics*, *20*(1), 156–167. https://doi.org/10.1093/bib/bbx086
36. Ienca, M., Ferretti, A., Hurst, S., Puhan, M., Lovis, C., & Vayena, E. (2018). Considerations for ethics review of big data health research: A scoping review. *PloS one*, *13*(10), e0204937. https://doi.org/10.1371/journal.pone.0204937
37. Panch, T., Mattie, H., & Atun, R. (2019). Artificial intelligence and algorithmic bias: implications for health systems. *Journal of global health*, *9*(2), 010318. https://doi.org/10.7189/jogh.09.020318
38. Baro, E., Degoul, S., Beuscart, R., & Chazard, E. (2015). Toward a Literature-Driven Definition of Big Data in Healthcare. *BioMed research international*, *2015*, 639021. https://doi.org/10.1155/2015/639021
39. Mooney, S. J., & Pejaver, V. (2018). Big Data in Public Health: Terminology, Machine Learning, and Privacy. *Annual review of public health*, *39*, 95–112. https://doi.org/10.1146/annurev-publhealth-040617-014208
40. Lai, C. C., Shih, T. P., Ko, W. C., Tang, H. J., & Hsueh, P. R. (2020). Severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) and coronavirus disease-2019 (COVID-19): The epidemic and the challenges. *International journal of antimicrobial agents*, *55*(3), 105924. https://doi.org/10.1016/j.ijantimicag.2020.105924
41. Ye, Q., Wang, B., Mao, J., Fu, J., Shang, S., Shu, Q., & Zhang, T. (2020). Epidemiological analysis of COVID-19 and practical experience from

China. *Journal of medical virology*, *92*(7), 755–769. https://doi.org/10.1002/jmv.25813

42. Kruse, C. S., Goswamy, R., Raval, Y., & Marawi, S. (2016). Challenges and Opportunities of Big Data in Health Care: A Systematic Review. *JMIR medical informatics*, *4*(4), e38. https://doi.org/10.2196/medinform.5359
43. Price, W. N., 2nd, & Cohen, I. G. (2019). Privacy in the age of medical big data. *Nature medicine*, *25*(1), 37–43. https://doi.org/10.1038/s41591-018-0272-7
44. Sun, K.; Chen, J.; Viboud, C. Early epidemiological analysis of the coronavirus disease 2019 outbreak based on crowdsourced data: A population-level observational study. Lancet Digit. Health 2020
45. Qin, L.; Sun, Q.; Wang, Y.; Wu, K.F.; Chen, M.; Shia, B.C.; Wu, S.Y. Prediction of Number of Cases of 2019 Novel Coronavirus (COVID-19) Using Social Media Search Index. Int. J. Environ. Res. Public Health 2020, 17, 2365
46. Yang, Z.; Zeng, Z.; Wang, K.; Wong, S.S.; Liang, W.; Zanin, M.; Liu, P.; Cao, X.; Gao, Z.; Mai, Z.; et al. Modified SEIR and AI prediction of the epidemics trend of COVID-19 in China under public health interventions. J. Thorac. Dis. 2020, 12, 165–174
47. Pinto, D., Park, Y., Beltramello, M. *et al.* Cross-neutralization of SARS-CoV-2 by a human monoclonal SARS-CoV antibody. *Nature* **583,** 290–295 (2020). https://doi.org/10.1038/s41586-020-2349-y
48. Song, J.; Song, T.M.; Seo, D.C.; Jin, D.L.; Kim, J.S. Social Big Data Analysis of Information Spread and Perceived Infection Risk During the 2015 Middle East Respiratory Syndrome Outbreak in South Korea. Cyberpsychol. Behav. Soc. Netw. 2017, 20, 22–29.
49. Pacheco Rocha, N.; Dias, A.; Santinha, G.; Rodrigues, M.; Queirós, A.; Rodrigues, C. Smart Cities and Healthcare: A Systematic Review. Technologies 2019, 7, 58.
50. 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 in the Populations when Cities/Towns Are Under Quarantine. Infect. Control Hosp. Epidemiol. 2020
51. Allam, Z.; Jones, D.S. On the Coronavirus (COVID-19) Outbreak and the Smart City Network: Universal Data Sharing Standards Coupled with Artificial

Intelligence (AI) to Benefit Urban Health Monitoring and Management. Healthcare 2020, 8, 46.

52. Hua, J.; Shaw, R. Corona Virus (COVID-19) "Infodemic" and Emerging Issues through a Data Lens: The Case of China. Int. J. Environ. Res. Public Health 2020, 17, 2309.
53. Perrella, A.; Carannante, N.; Berretta, M.; Rinaldi, M.; Maturo, N.; Rinaldi, L. Novel Coronavirus 2019 (Sars-CoV2): A global emergency that needs new approaches? Eur. Rev. Med. Pharmacol. Sci. 2020, 24, 2162–2164
54. Wong, Z.S.Y.; Zhou, J.; Zhang, Q. Artificial Intelligence for infectious disease Big Data Analytics. Infect. Dis. Health 2019, 24, 44–48.
55. Aiello, M., Cavaliere, C., D'Albore, A., & Salvatore, M. (2019). The Challenges of Diagnostic Imaging in the Era of Big Data. *Journal of clinical medicine*, *8*(3), 316. https://doi.org/10.3390/jcm8030316
56. Wu, P. Y., Cheng, C. W., Kaddi, C. D., Venugopalan, J., Hoffman, R., & Wang, M. D. (2017). -Omic and Electronic Health Record Big Data Analytics for Precision Medicine. *IEEE transactions on bio-medical engineering*, *64*(2), 263–273. https://doi.org/10.1109/TBME.2016.2573285
57. Agrawal, R., & Prabakaran, S. (2020). Big data in digital healthcare: lessons learnt and recommendations for general practice. *Heredity*, *124*(4), 525–534. https://doi.org/10.1038/s41437-020-0303-2
58. Luo, J., Wu, M., Gopukumar, D., & Zhao, Y. (2016). Big Data Application in Biomedical Research and Health Care: A Literature Review. *Biomedical informatics insights*, *8*, 1–10. https://doi.org/10.4137/BII.S31559
59. Sanchez-Pinto, L. N., Luo, Y., & Churpek, M. M. (2018). Big Data and Data Science in Critical Care. *Chest*, *154*(5), 1239–1248. https://doi.org/10.1016/j.chest.2018.04.037
60. Peng, F., Tu, L., Yang, Y., Hu, P., Wang, R., Hu, Q., Cao, F., Jiang, T., Sun, J., Xu, G., & Chang, C. (2020). Management and Treatment of COVID-19: The Chinese Experience. *The Canadian journal of cardiology*, *36*(6), 915–930. https://doi.org/10.1016/j.cjca.2020.04.010
61. Chen, C. M., Jyan, H. W., Chien, S. C., Jen, H. H., Hsu, C. Y., Lee, P. C., Lee, C. F., Yang, Y. T., Chen, M. Y., Chen, L. S., Chen, H. H., & Chan, C. C. (2020). Containing COVID-19 Among 627,386 Persons in Contact With the Diamond Princess Cruise Ship Passengers Who Disembarked in Taiwan: Big Data

Analytics. *Journal of medical Internet research*, *22*(5), e19540. https://doi.org/10.2196/19540

62. Chen, F. M., Feng, M. C., Chen, T. C., Hsieh, M. H., Kuo, S. H., Chang, H. L., Yang, C. J., & Chen, Y. H. (2020). Big data integration and analytics to prevent a potential hospital outbreak of COVID-19 in Taiwan. *Journal of microbiology, immunology, and infection = Wei mian yu gan ran za zhi*, S1684-1182(20)30104-3. Advance online publication. https://doi.org/10.1016/j.jmii.2020.04.010

Part 2- Online Course Report

Name of the Course-Improving Global Health: Focusing on Quality and Safety.

COURSE DESCRIPTION

Achieving the highest standards of quality in healthcare is a universal problem, and all the healthcare professionals, doctors and nurses are facing issues in providing effective care. Thus, this online course was launched by Harvard Initiative on Global Health Quality (HIGHQ)– Improving Global Health: Focusing on Quality and Safety.

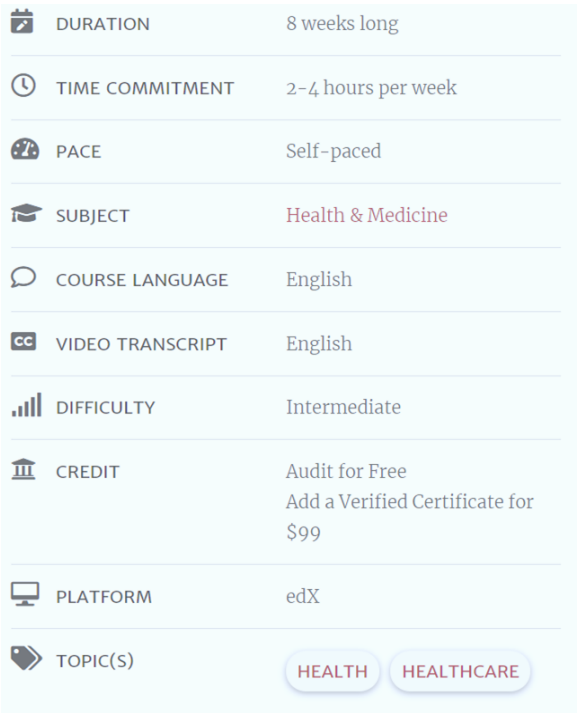
It was an eight-week course which comprised of eight modules focussing on different objectives. It serves the purpose of learning the measurement of quality and how it can be improved.

Improving the accessibility and quality of healthcare has a growing importance in today's world. This course recognizes the ideas and ways of transformation of safety and quality standards in health systems. It aims at strengthening the collective understanding of patient-centred safe quality care to deliver patient satisfaction.

Each section of this course offers interactive platforms including surveys, graded quizzes, tools, question-answers windows and feedbacks.

This course covers a broad range of topics including the importance of quality, measurement of quality, improvement of quality and safety and various case studies. It offers a highly comprehensive and interactive curriculum.

The following image shows the course description as mentioned on the website:



OBJECTIVES

This course comprises of 8 different lessons. Each lesson aims on different objectives. Here are some objectives from each of the 8 lessons.

- BURDEN- Understanding the basic terms, measurement of quality in different countries, prevalence of tuberculosis and identification of strategies and innovations to improve the treatment of tuberculosis
- MANAGEMENT- To understand the role of managers in quality control, measurement of quality management, example of Virginia mason Hospital system and JIPMER in improving quality
- IT & DATA- To understand health Information exchange, IT in data management and utility of electronic health records
- IMPROVEMENT- To understand the framework for quality standards, characteristics for healthcare leadership and accountability in health care quality.
- STANDARDS- To Understand the role of legal regulations in quality control, factors impacting healthcare quality, discuss practitioner licensing and patient compensation
- MEASUREMENT- To Understand quality variation across health systems globally, categories of health resources and allocation, global health quality indicators, comparison of countries based on measurement of quality.
- PUBLIC SYSTEMS- To understand the growth of quality improvement, eras and categories of quality movement, in healthcare. To discuss health care reform in Mexico and to determine the role of WHO in quality care
- PATIENTS- To understand patient-centred care, patient experience and improve outcome, impacts of social media on health care providers, features of high quality care and reasons for low quality, women delivery preferences determining quality.

COURSE CONTENTS

The following table (Table 1.) depicts the contents of the online course, along with sub-parts of each lesson:

Table 1. Course Contents

S. NO.	LESSONS	SUB-PARTS	TAUGHT BY
1.	Burden	• Burden And Harm • Tuberculosis	David Bates Madhukar Pai
2.	Measurement	• Quality Variation • Global Metrics	Jishnu Das Niek Klazinga
3.	Standards	• Legal System and Regulations	Allen Kachalia
4.	Improvement	• Quality Improvement • Case of Rwanda	Maureen Bisognano Agnes Binagwaho
5.	IT and Data	• Health Information Exchange	Julia Adler-Milstein
6.	Management	• Role of Management • Case of India	Raffaella sadun T.S. Ravikumar
7.	Patients	• Patient Centred care • Maternal and new born care • New technology	Ronen Rozenblum Margaret Kruk Felix Greaves
8.	Public Systems	• The Quality Movement • National Strategies and Case of Mexico • Role of World Health organization	Don Berwick Julio Frenk Ed Kelley

SUMMARY OF EACH LESSON

The following table (table 2.) comprises of the summary of each lesson in the online course:

Table 2. Summary

S. No	Lesson	Summary
1.	Burden	This lesson comprises of two sub-parts: Burden and harm and Tuberculosis. In this lesson Burden and harm in patient care, the preventability of harm, comparison of High and low income countries in terms of patient safety, ambulatory care and medication errors are being discussed. Apart from this a detailed background about tuberculosis, it's treatment by public and private sector and post diagnostic monitoring, is provided.
2.	Measurement	In this lesson, quality variation and Global metrics have been discussed, including resources allocation and variation and the need for advanced data infrastructure to maintain quality and safety globally.
3.	Standards	In this section a comprehensive understanding on the legal and regulatory frameworks acting in health systems has been provided. It includes accreditation, staffing ratios, patient compensation and transparency
4.	Improvement	This lesson focuses on quality improvement in health systems and the steps for improvement. Plan-Do-Study-Act is also mentioned. The second sub-part focuses on the Case study of Rwanda.
5.	IT and Data	This lesson includes lectures on Health information exchange-background and applications, and implementation on a global level.

6.	Management	Role of Management in ensuring patient safety and quality of care delivered is mentioned in this lesson. It also included the Case of India- JIPMER and the launch its quality control cell.
7.	Patients	This section aims to provide a deep understanding on patient centred care, measuring experience, ambulatory care and the improvement of overall process. Maternal and new born care delivery is also taught along with latest technologies in online feedback tools to deliver patient satisfaction.
8.	Public Systems	This lesson delivered lectures on quality movement in healthcare along with Era 3 and triple aim of healthcare. It also provides information on framing global health quality and Case study of Mexico. The course ended with a lecture on the Role of WHO in ensuring high standards of quality and safety of patients worldwide.

LEARNING METHODOLOGY

The Online course Improving Global Health: Focussing on Safety and Quality, allows students to view lectures, access materials, and collaborate with peers on their own schedule. Lectures are pre-recorded and presented with instructor as well as voice-over. This delivery methods allow students to review and re-review lessons as necessary. This option of pre-recorded lecture is extremely helpful could be quite as it allows the students to work through lessons at their own convenience.

This course employs the following learning methodologies:

- High quality downloadable pre-recorded lectures
- Microsoft PowerPoint presentations with and without voice-over
- Forums and discussion boards
- Email communication
- Google Drive and similar collaborative tools
- Tools for off-hour support
- Suggested readings which include Research publications which appear at the end of each session and aims to provide in-dept knowledge
- Survey tools for course feedback

KEY LEARNINGS

Improving the quality of healthcare delivery and Safety of Patients has become a global priority. However, achieving the highest levels of quality in healthcare and ensuring maximum patient safety is a difficult task especially in countries with low-and-middle income (LMICs). In the view to achieve this target, the Harvard Initiative on Global Health Quality (HIGHQ) launched a massive open online course (MOOC) on the edX platform which offers in depth understanding on patient safety and quality of care. This course aims to give an idea on the innovative educational partnership between the course team and one of these countries, Rwanda. This course also includes case of India in handling Tuberculosis and a case study of Mexico.

This course focusses on specific ways to improve healthcare quality by measuring and comparing the quality and safety in China and India. The quality is measured in terms of variation which includes, public-private variation and inconsistent allocation of resources. It also identified measurements of quality in ambulatory care.

The safety of the patients are compromised due to the lack of diligence of doctors and nursing stuff. The commonly caused infections are usually preventable, still the patients are at a risk. To prevent these infections some points can be kept in mind like hygiene practices by doctors, adhering to the protocols and profiling people into risk categories.

Distinct barriers to accessing and improving quality care and patient safety has been identified in LMICs. Thus it is important to improve the health quality and safety in such countries. In this course a case study on Rwanda, East Africa was analysed to find out that Rwanda's health sector emerged glorious after being completed destroyed in 1994. As a country they achieved the Millennium Development Goals on healthcare, which included maternal health, new born health and HIV/AIDS.

This course also threw some light on the technological innovations and use of data science in Health Information Exchange, how health data is stored and analysed globally. The role of management plays a vital role in ensuring health safety and patient quality. Efficient and effective management has a direct impact on quality care and patient safety

In addition to Rwanda's health care model, this course included the case of JIPMER, India, a leading establishment. The particular section focussed on how JIPMER launched the Quality Council and it's working strategy to ensure Quality.

Lastly, it focussed on Patient centred care and Public Systems including ways of measurement and improvement of patient centred care and the role of World Health Organization in improving the quality and overall health outcomes of the health systems around the world.

CONCLUSION

The online course Improving Global Health: Focussing on Safety and Quality was successfully enrolled and completed. As a management student, it helped to expand exposure to revelant knowledge and skills which focussed on improving healthcare delivery. This course provided the desired outcome in helping to understand the working of quality cells within organizations of different countries. It provided practical based case studies and examples for improving the quality and patient safety. The suggested reading material along with every lesson helped in gaining in-dept knowledge and cleared doubts simulataneously.

This course comprised of engaging and interactive lessons and was well aligned with the current health sector priorities to build capacity in improvement of quality and safety among healthcare professionals.

This course is freely accessible to everyone and does not require any pre-requisite which adds to the many other benefits.