Big Dataanalytics inthe medical caresector: Sources, challenges, applications, future prospects

Atirek Raj Singh[a],Kuldip Katiyar[b]
Department of Mathematics,
Chandigarh University,
Gharuan, Punjab, India,
[a]rajatirek2@gmail.com,
[b]kuldeepe7279@cumail.in

Abstract :

In recent years, various institutions around the globe have produced abundantly haphazard data, and this unorganized data is called Big data. These are not only large in size but also diverse and rapid which is a strenuous task to manage with existing tools and processes. Electronic health records (EHRs) or medical data collected from hospitals or patients are among the fastest-growing types of data. Owing to the rapid rise of medical data, it is necessary to develop appropriate tools and approaches for processing and extracting values and knowledge from these datasets to upgrade patient care and reduce medical costs. The medical care industry faces the challenge of managing profuse info from different sources. Although Various technologies and strategies have been developed in healthcare institutions to process these vast amounts of data but it is still an undergoing process. The idea behind this review paper is to give an overview of the impact of big data analysis in medical care. At first, we define and elucidate the characteristics and sources of these data in medical care. After that, we shall go over its challenges in implementing this, and applications related to health management, and finally, wedraw conclusions.

***Key Words*:** Big data; Medical care; Electronic Health Records(EHRs); EMR

I. Introduction:

Big Data is rapidly being prompted in all industries, andthe health sector is no different. As the world's population grows withlongevity, treatments are changing rapidly, and most of the decisions that drive rapid changes need to be data-based.

Pharmaceutical industry experts and stakeholders have already started to analyze big data on a regular basis to generate insights, but their efforts are still in their infancy and must be integratedto confront healthcare issues and enhance quality.Collecting, storing, and retrieving data are all parts of medical informatics that helpmedical care professionals achieve better results. By linking a wide variety of biomedical data sources such as sensor data, images, gene sequences, clinical tests, and demographics, medical statistics is characterized by its heterogeneity and diversity. Most records in the medical system (such as Physician notes, laboratory test results, and clinical reports are not structured and electronically stored, so that means this is solely available in paper format and is constantly increasing. Currently, the focus is on digitizing these vast collections of paper. To achieve this goal, a data size revolution poses a problem. The five Vs of volume, diversity, and speed are addressed by numerous terms and models that have emerged to solvecomplications related to big data. EHRs (ElectronicHealthRecords), sensor details, health information exchanges, patient enrollments, hereditary databases, and public case studies areexamples of data types used in health applications. In the medical industry, public records are a vital source of informationand require effective data analysis to address the challenges they create.

A. Big Data:

Big data expresses large, hard-to-manage volumes of data,be it structured or unstructured. It is a field to analyze, and systematically extract information for dealing with data sets that are too giganticand

eISSN 1303-5150

complicated todeal with by existingdata-processing applications.

Even though several definitions of big data already exist, but the most sought-after and well-acknowledged definition is given as 5 V's**[1].**

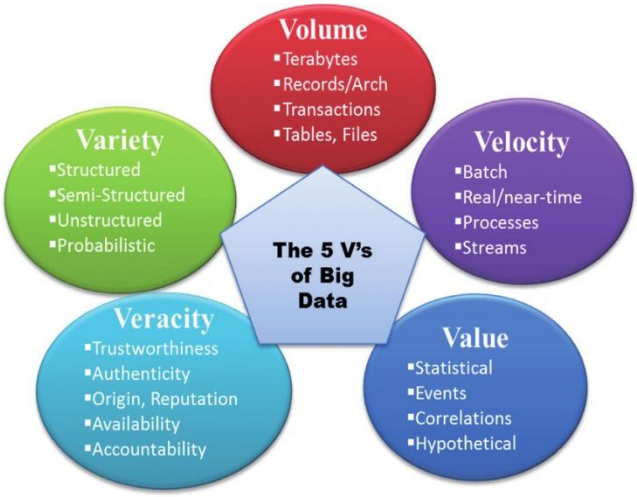

Figure 1: Characteristics of big data

ImageSource: researchgate.net/figure/Characteristics-of-big-data-or-5V_fig1_339897935 [2]

- **Velocity:** It refers to the speed atwhich data is being induced. These data can be generated fromvarious sources like machines, industrial networks, social media, mobile phones,and many more.
- **Volume**: Itsaysabout the size of data and whether the data can be considered Big data or not, based on the volume.
- **Variety:** Any data that is structured, unstructured, or semi-structured, classified under the category of variety.
- **Veracity:** As we collect the pieces of information from various sources, we first have to check for accuracy before using them for our useful insights, whichassures us of quality or integrity.
- **Value:** Any data which cannotprovide us an idea about something productive is of no use irrespective of the fact how much info we collected from different sources. We ought to extract the value out of data.

II. *Sources of Medical-Care Data:*

B. Big Data Characteristics:

Data captured, collected, and accumulated within the medical fieldmay be organized and disseminated, taken from different origins, and in different configurations and formats. This includes facts and figures on physiological,clinical, and environmental disorders, medical images, disease management, drug prescriptions, nutritional parameters **[3],** and many more**.** Many different studies agree on a typical source of healthcare data of which some of them are discussed below:-

- **Electronic Health Records (EHRs):**This is an electronic record of patients'case history, maintained by the service provider over a period of time. EHR may contain recordsof the clinical and management meetings between care providers (doctors, nurses, etc.) and patients **[4].**It includes various data such as demographics, medical background, medications, allergies, vaccination status, laboratory outcomes, radiographic images, personal info such as BMI, billing information, active medical issues, etc. **[5]**.
- **Electronic Medical Record (EMR):** EMR is similar to EHR and is a digital record of patient information maintained in the pattern of a chart, includingpatient's medical and treatment records from the clinic. Normally, the digital file remains in the clinicand is not dispensed even if patients change doctors, their EMR is not shared. Although, this record chart is preserved in the database of health centersorhospitals**[6].**
- **Patient-ReportedOutcomes (PRO**): Described asdirect patient reports onpatients' healthand treatmentbased on theirsagacity of the disease and its cure. This infocomprises several endpoints like symptoms, health status, and quality of life **[7].**
- **wearable Device**: Most wearable devices can collect biochemical, physiological, and motion-capture informationlike (human heart rate,steps, and blood pressure). Therefore, it collects health info from patients and has data-sharing capabilities **[8].** Analysis of allsorts of information, when integrated into EHRs, is beneficial for keeping a watch on and diagnosing the health of the assorted chronic state.
- **Data collected fromsocialplatforms:** Patients' posts on social platformslike Facebook, Instagram, Twitter, etc. can be oftendrawn out to

eISSN 1303-5150

get information about disease trends, their satisfaction, happiness, andinterests**[9].**

III.Challenges related tomedical care data:

Obviously, Big data has a huge impact on Healthcaremanagement but at the same time, it is not that easy to implement in this particular field.Here are some of the challenges that we might face:-

- **Aggregation of Data**: Patients' data are often distributed across manyhospitals and servers. Collecting it all together, a lot of planning is needed to ensure that all data creators work together as new data is created in the future. Moreover, each engaged organization must understand and agree on the type and format of data they want to evaluate. Aside from the issue of storage formats (paper, film, traditional databases, EHR, etc.), it is essential to establish the accuracy and quality of such figures. This requires a review of data governance aswell as data cleansing (primarily a manual process).
- **Management**: Ultimately, organizations need to adapt their business practices to realize the latent of Data analytics in the medical sector. Data scientists may need the assistance of IT staff who have the ample skills to perform the scrutiny of data. Some of the institutions may struggle with the plans that their IT infrastructure needs to be "ripped and replaced," but cloud service providers reduce some of these concerns to some extent. Managers and physicians can take some time to depend on the previously unseen advice that big data can offer.
- **Challenges of Policy and Process:** Once your data is substantiated and collected, you need to address various process and policy-related concerns. Health Insurance Portability and Accountability Act (HIPAA) regulations requisite policies and procedures to protect health-relatedinformation. Access control, authentication, security in transit, and other rules may complicate the task,and these multifaceted problems have been resolved somewhat bycloud service providers, prominently, AWS of Amazon, which provides HIPAA and Protected Health Information (PHI) compliant cloud services.
- **Storage**: Storing enormous amounts of data is the biggest challenge in itself, but many companies are accustomed to storing data on their premises which gives them an edge in terms of security, access, and uptime control. Still, scaling your on-premises server network can be exorbitant and hard to maintain. Cloud-based storage withinfrastructure seems to be a preferablechoice for many healthcare organizations as costs decrease and reliability increases. Organizations need to choose cloud partners who understand the seriousness of health-related compliance and security issues. In addition, cloud storage provides reduced upfront costs, flexible disaster recovery, and ease of extension**[10].** This could be the foremost flexible and viable approach for vendors with different data access and storage need.
- **Cleaning**: First, Data need to be cleaned to ensure post-collection authenticity, precision, constancy, relevance, and naturalness. This cleaning process will be automated manually or using logical rules to make sure a high level of precision and integrity. Some of the trailblazing and accurate tools use machine learning technology to avoid wasting time,and money and to prevent unimportant data from destroying actual data projects.
- **Precision:**Many studies have noticed that delineating patient's data on EMR or EHR is not yet completely error-free**[11-14].** All of these aspects can affectits quality throughout the life cycle. We can improve the quality of documentation by making use of patients' self-proclaimed questionnaires about symptoms.
- **Security:** With a number of security breaches, and hacker attacks, the safety of data is a topmost priority for medical care institutions. After identifying some vulnerabilities, an inventory of technical safeguards within the Protected Health Information (PHI) was formed. These ordinances of HIPAA help organizations use storage, outbound, and authentication protocols to regulate access, integrity, and auditing. Many problems can be avoided with common security measures such as antivirus software, firewalls, encryption of delicate data, and multi-factor validation.
- **DataDispensing**: Exchanging data with other medical institutions is essential and usual but if data cannot interoperate during such sharing,data movement between different institutions can be rigorously confined. This may be because of technical and institutional barriers. It can lead

eISSN 1303-5150

doctors to lose important details for making conclusions about patient follow-up and treatment approaches. Some Solutions such as "Fast Healthcare Interoperability Resource (FHIR)", "publicAPIs",and "Carequality" (aninteroperability framework) make interoperability and information sharing simple and confidential. The largest barrier to data sharing is treating data as a product that can produce a competitive advantage. Therefore, sometimes both vendors and providers may deliberately block the flow of information amongst different EHRs systems **[15].**

Integration of healthcare data, including individual patient data and epigenomics, has been a major challenge. Data warehousing, link integration, service-oriented architectures, view integration, and merging data from several Web-based resources for a new Web application are just a few of the many various integration approaches that have been developed. In addition to the difficulties outlined above, data fusion involves other important considerations.

Health providers need to conquer all the challenges of this list to instigatea Big Data exchange environmentwhichoffers reliable, timely, and meaningful info by joining the care continuum.Overcoming these challenges requires time, commitment, funding,and communication.

IV. Application:

In this part, we are going to discuss the utility of information in the medical field. Here are a numberof key areas where big data can assist the healthcare systems –

- **Recognizing high-risk patients**: Patients in danger of chronic disease make up the majority of the patient population. It's imperative that the hospital identify them early and supply an aggressive treatment plan to make sure that the condition is stable and does not get worsen. Advanced analytical tools can be implemented by physicians to distinguish factors that influence risk andbased on this, define treatment and intervention plans correspondingly**[16].**
- **Elimination of mediation mistakes:**selecting the wrong medicine can harm the patient's well-being and is a major cause of patient death. Big data helps analyze the records of patients taking prescription drugs and helps reduce error rates by flagging potential nonstandard errors**.[17]**
- **Avoid fraud and security breaches:** Health maintenance-relatedinfobreaches can prove to be costly as well as potentially fatal as hackers can easily get access to delicate medical info associated with the patients. The latest research shows that the medical sector is at a 200% increased risk of information breaches because the info is so valuable. With this in mind, healthcare institutions using big data have been able to stay away from fraud, security breaches, and threats.
- **Provide remote monitoring function:** Analytics of big data offers many opportunities for remote monitoring capabilities via biometric devices and mobile data acquisition applications to play an important part in observing a patient's blood pressure and blood sugar levels, medications, and activity levels. These data can be efficiently utilized by care providers to handle patients with chronic illness, and such remote monitoring support enhances their involvement and provides a quality look after for patients at home or in the hospital**.[18]**
- **Reduce patient re-admissions and needless complications:** Big data and its analysis can reduce the re-admission rate of hospital patients. Predictive analytics can help classify discharged patients based on readmission risk. It also can play a prominent role in examining the effectiveness of the various types of interventions implemented for re-admissions. Moreover, big data technology can be used to reduce the likelihood of complications, especially if the patient is hospitalized at the time of infection.
- **Better health administration of the population**: Medical care networks are increasingly bringing about new programs to manage the good state of health of the population and promote patient involvement. By utilizing analytics in healthcare management, healthcare providers can effectively validate problems through EMR and identify high-risk patients within the target population by diagnosing and analyzing test results.Appropriate diagnostic methods must be employed before assessing the severity of diseases.
- **Support for Patient-CenteredMedical Home(PCMH):**It isa care model, that ensures that primary care physicians coordinate patient care and receive the care they need, when and where

eISSN 1303-5150

they need it. Many caregivers rely on analytics and databases to provide patient-centric healthcare facilities. For Example – Analyzing EMR data can lead to interventions by healthcare providers with regular monitoring in diabetic patients with potentially high HbA1c levels**.[19]**

V.Future Prospects:

The healthcare sector is booming rapidly, and the need to manage patient care and innovative care is increasing as a synonym. With this growing demand, new technologies are being introduced into the industry. One such major change that may occur in the time ahead is the practice of "Big Data" and its analytics in medical care.

According to a report from the International Data Corporation (IDC), *"healthcare statistics aresupposed to grow at a compound annual growth rate (CAGR) of 36% by 2025"***.** This is quicker than other industries like manufacturing, financial services, and media. There are several ways big data can help and change the entire healthcare scenario:

- **Health Tracking**

  With the employment of massive Data and its analytics within the medical care sector, healthcare competent can now trace the vitals and details of various patients and service users. This is often pioneering how effective medical care is going to be because it keeps a check on patients in an exceedingly manner that has never been done earlier. Through health tracking, not only can a patient's pulse, sleep, and exercise be monitored, but recent medical innovations can now go that one step ahead and record a patient's pulse, pressure, and sugarlevels**[20].**It enables a more proactive and dynamic-thinking medical care institution, as health professionals can curb those who spend time within the hospital by tracking and intervening before a significant pathological state takes place.

- **Cost Reducing**

  The expenses of medical care don't come economically, and plenty of hospitals find themselves combating the assorted costs that come from offering medical treatment. Therefore, Big Data having the ability to cut back expenses could be a desirable addition to the medical care industry. It can be practicable as Big Data can utilize predictive analysis to anticipate admission rates, which might help with the allotment of staff. This assists the hospital to understand what percentage of staff members are required and stops them from under or overbooking, saving them capital and resources**[21].**

  Using health tracking, and checking a patient's health before hospitalization is required, this could lessenfatal health problems, and also the capital that must be spent on treatment for patients. This may amount to substantial savings, and research has proposed that predictive analysis could save medical institutions approximately 25 percent of their annual costs over the comingyears.

- **Assisting Patients at Higher Risk**

  High-risk patients can now be better assisted and monitored thanks to the digitization of hospital and patient records. If a patient has been admitted to the hospital several times, for example, his health records can be obtained to uncover trends in his health ailment and any chronic difficulties. This will have a significant impression on today's medical care system, as this advancement in knowledge will allow doctors to deliver better care to patients and lessen the numeroushospital visits that patients must endure and hospitals' bills.

- **Human Errors are Reduced**

  Human errors have resulted in many instances when imprecise prescription has been prescribed or appointments have been missed, both of which can have serious repercussions. While certain human errors are unavoidable when dealing with big amounts of data, Big Data will help to limit the number of errors produced. The software can track data from various medical experts about patients and flag up any prescription errors, potentially saving lives. This program is ideal for healthcare practitioners that see a large number of patients in a single day.

- **Improvements in the Medical care Institution**

  The practice of hugedata and analytics in the healthcare field opens the door to a variety of new professions and jobs. However many have been skepticalaboutthis big data's impact on conventional care and therapy models, the digitization of medical information and usage of software tools enables staff to spend more time treating patients and less time recording. Correspondingly, the information and data provided by these innovative solutions can now

eISSN 1303-5150

benefit employees.There are also new job opportunities that have arisen as a result of the employment of hugedata, like DNP to ENL programs, which prepare nurses with executive knowledge and competencies in data-driven approach, leadership skills, and also the potential to collaborate with and establish transformative care models. This is not simply a road for advancement, but it also allows employees to cash in on a premium DNP executive leadership compensation**[22].**

- **Increase in Patient Engagement**

With the practice of Big Data and new treatment models, patients can take account of their own health and well-being which might push them to be more health-consciousand lead a healthier lifestyle. According to research,the market for smart devices has boomed in recent years, resulting in greater interest than ever before in tracking one's heart rate, steps taken, and sleeping patterns. Without the presence of a medical expert, all of this traceable and vital data can be captured on a regular basis using these devices. This allows people to take a more active role in their own well-being by keeping track of changes and being aware of any emerging issues related to health.This will have a significant influence on medical professionals' and organizations' time and resources**.**

For some, the viewpoint aboutbigdata playing a significant part in medical care is still in vague, but it has the potential to remould how we receive medical treatment, the standard of that treatment, and the resources dedicated to making a better way of living.

VI. Conclusion:

I have presented an extensive definition and influence of big data in medical careadministration in this paper. Big data plays a significant role in healthcare informatics and has a substantial impact on the healthcare system by allowingdoctorstoprescribe suitable pharmaceuticals for each individual rather than those that work for the majority of individuals.The fusion of big data and medical care analytics can lead to therapies that are beneficial for particular patients. Big data has already started to revolutionize andtransformthe healthcare sector ahead in exceedinglydiverseways. In medical care, itassuresadvancement in medicine, technology, and financing resulting in solutions that upgrade patient care and increase value in medical care institutions.However, this will necessitate collaboration and innovation amongst stakeholders along with providers, pharmaceutical producers, government and policymakers, and scientific and research communities, so as to rethink the outline and functioning ofsystems. They need to provide the technology infrastructure needed to contain and merge the vast amount of medical care data which is expected to multiply, according to field evaluation. They must also back human capital, including Information Technologyconsultants, researchers, databasearchitects, and databaseadministrators, in order to lead us into this creative and inspiring frontier of the well-being of humans.

References

[1] Ishwarappa, & Anuradha, J. (2015). "A Brief Introduction on Big Data 5Vs Characteristics and Hadoop Technology." Procedia Computer Science, 48, 319–324.

[2]https://www.researchgate.net/figure/Characteristics-of-big-data-or-5V_fig1_339897935

[3] S. G. Alonso, I. Torre-Diez, J. Rodigues, S. Hamorioui, and M. L. Coronado, (2017) "A Systematic Review of Techniques and Sources of Big Data in the Healthcare Sector.", Journal of Medical Systems, Vol. 41, No. 11, pp.183

[4] A. P. Ambinder, (2005) "Electronic Health Records By", Journal of Oncology Practice, Vol. 1, No. 2, pp. 57–63

[5] R. S. Evans, (2016) "Electronic Health Records: Then, Now, and in the Future", Yearbook of medical informatics, pp. S48–S61.

[6] Z. Liang et al., (2014) "Deep learning for healthcare decision making with EMRs", Proceedings - 2014 IEEE International Conference on Bioinformatics and Biomedicine, IEEE BIBM 2014, (Cm), pp. 556– 559

[7] M. Calvert et al., (2015) "Putting patient-reported outcomes on the Big Data Road Map", Journal of the Royal Society of Medicine, Vol.108, No.8, pp. 299–303

[8] M. Uddin, and S. Syed- abdul, (2020) "Data analytics and applications of the wearable sensors in healthcare: An overview", Sensors (Switzerland), Vol.20, No.5

eISSN 1303-5150

[9] J. Zhang, N. Xue, and X. Huang, (2016) "A Secure System for Pervasive Social Network-Based Healthcare", IEEE Access, Vol.4, pp. 9239–9250

[10] Dash, S., Shakyawar, S.K., Sharma, M. et al. Big data in healthcare: management, analysis and future prospects. J Big Data 6, 54 (2019)

[11] Valikodath NG, et al. Agreement of ocular symptom reporting between patient-reported outcomes and medical records. JAMA Ophthalmol. 2017;135(3):225–231.

[12] Fromme EK, et al. How accurate is clinician reporting of chemotherapy adverse effects? A comparison with patient-

reported symptoms from the Quality-of-Life Questionnaire C30. J Clin Oncol. 2004;22(17):3485–3490

[13] Beckles GL, et al. Agreement between self-reports and medical records was only fair in a cross-sectional study of the performance of annual eye examinations among adults with diabetes in managed care. Med Care. 2007;45(9):876–883

[14] Echaiz JF, et al. Low correlation between self-report and medical record documentation of urinary tract infection symptoms. Am J Infect Control. 2015;43(9):983–986.

[15] Adler-Milstein J, Pfeifer E. Information blocking: is it occurring and what policy strategies can address it? Milbank Q. 2017;95(1):117–135.

[16] Bates, D. W., Saria, S., Ohno-Machado, L., Shah, A., & Escobar, G. (2014). "Big Data In Health Care: Using Analytics To Identify And Manage High-Risk And High-Cost Patients." Health Affairs, 33(7), 1123–1131

[17]Chaiken, B. P., &Holmquest, D. L. (2003). Patient safety: modifying processes to eliminate medical errors. Nursing Outlook, 51(3), S21–S24.

[18]María Luz Morales-Botello Diego Gachet Manuel de Buenaga Fernando Aparicio María J Busto. "Chronic patient remote monitoring through the application of big data and internet of things." 2021Volume: 27(3)

[19]Bates, David W., and Asaf Bitton. "The future of health information technology in the patient-centered medical home." Health affairs 29.4 (2010): 614-621.

[20] Dimitrov, Dimiter V. "Medical internet of things and big data in healthcare." Healthcare informatics research 22.3 (2016): 156-163.

[21]Wang, Yichuan, and Nick Hajli. "Exploring the path to big data analytics success in healthcare." Journal of Business Research 70 (2017): 287-299.

[22]Srinivasan, Uma, and BavaniArunasalam. "Leveraging big data analytics to reduce healthcare costs." IT professional 15, no. 6 (2013): 21-28.

eISSN 1303-5150