www.ijprems.com
editor@ijprems.com

INTERNATIONAL JOURNAL OF PROGRESSIVE RESEARCH IN ENGINEERING MANAGEMENT AND SCIENCE (IJPREMS)
(Int Peer Reviewed Journal)
Vol. 04, Issue 09, September 2024, pp : 400-404

e-ISSN : 2583-1062
Impact Factor : 5.725

A COMPREHENSIVE REVIEW ON "THE SIGNIFICANCE OF BIG DATA ANALYTICS IN THE PHARMACEUTICAL AND HEALTH CARE SECTOR

P. Shailaja[1], P. Raga Sushmitha[2]

[1]Associate Professor, Department of Pharmaceutical Regulatory Affairs, Andhra University College of Pharmaceutical Sciences, Visakhapatnam- 530003, Andhra Pradesh, India.

[2]Student. Department of Pharmaceutical Regulatory Affairs, Andhra University College of Pharmaceutical Sciences, Visakhapatnam- 530003, Andhra Pradesh, India.

ABSTRACT

The health care and pharmaceutical industries are complex sectors in which a great amount of data from numerous sources is generated at a rapid pace. Big Data Analytics is a crucial trend in the pharmaceutical and healthcare industries for drug discovery, precision medicine, clinical trials, research and development, case-specific treatment, preventive health, electronic health records, retinal scans, biomarker data, consumer reviews, and more. This article discusses the potential for big data to revolutionize the health care and pharmaceutical sectors, as well as the significant benefits and challenges encountered in the development of data analytics; its impact on skill enhancement, the provision of patient-specific treatments, and the development of technology used.

Key words: Big data analytics, Precision medicine, Preventive health, Clinical trials.

1. INTRODUCTION

Significant quantities of data are extracted from many sources by the pharmaceutical and health care organizations with the primary goal of using this big data to support, improve, and implement better therapies for patients. Outdated are traditional data management practices including keeping hard copy records and reporting. The health care and pharmaceutical businesses must embrace digitization in the form of big data in this new era of rapidly increasing data generation in order to have a comprehensive understanding of public health. Big data is defined as "large volumes of high velocity, complex, and variable data that require advanced techniques and technologies to enable the capture, storage, distribution, management, and analysis of the data" in a study submitted to the US Congress in 2012. [1] Big data analytics is the process of gathering enormous amounts of unstructured data from a variety of sources in the pharmaceutical and healthcare sectors, including clinics, hospitals, regulatory agencies, medical records, patient care, Electronic Health Records (EHR), social media, demographics, genetic data, medical history, clinical reports, prescription drugs, and diagnostic reports. The data is then analyzed and interpreted to protect the public's health. Every day, the gathered data is combined with the aid of cutting-edge methods. Utilizing big data effectively in the pharmaceutical and healthcare industries aids in overcoming significant obstacles including the rise in healthcare expenses. Additionally, it promotes clinical improvement, ascertains and implements appropriate patient treatment protocols, guarantees patient safety, and effectively monitors the health care system. Big data in the healthcare sector facilitates a variety of data management tasks, such as tracking patient health and disease surveillance.

2. THE FIVE 'V'S OF BIG DATA

The five main features of big data, or the five V's, are as follows: volume, value, variety, veracity, and velocity.

1. Volume: The main big data traitor is this. It is the volume of information that pharmaceutical firms have examined and deciphered. Businesses generate petabytes or terabytes of data in order to do business and adhere to legal requirements.

2. Value: Better results, smarter strategic decision-making, and increased corporate efficiencies are how big data analytics adds value.

3. Variety: This is the main feature of big data. Every day, various forms of data such as those produced by electronic health records, medical devices, unstructured and structured data, clinical and non-clinical data, spreadsheets, audio, video, and images are removed from health care organizations. These data must be standardized. Three types of data exist: semi-structured (tagged and marked data without a formal structure like a database); unstructured (i.e., unorganized data); and structured (conventional databases, like SQL).

4. Veracity: The produced data must be precise and of the highest standard. It is established based on the acquired data's level of trust.

www.ijprems.com
editor@ijprems.com

INTERNATIONAL JOURNAL OF PROGRESSIVE RESEARCH IN ENGINEERING MANAGEMENT AND SCIENCE (IJPREMS)
(Int Peer Reviewed Journal)
Vol. 04, Issue 09, September 2024, pp : 400-404

e-ISSN : 2583-1062
Impact Factor : 5.725

5. Velocity: This refers to the data production pace. Rapid integration, analysis, evaluation, digestion, and reprocessing of the data is required. Data gathered from hospital patient monitoring devices in health care organizations gets to its destination swiftly.

SOURCES FORBIG DATA [3]

a) EHR: measures of height, weight, blood pressure, glucose, and pulse oximeter readings

b) Clinical decision support systems: prescription from a doctor; diagnostic reports; imaging reports; medical reports; records; patient medical history; demographics; genomic data; and information from medical insurance companies. c)social media: information from Facebook posts, blogs, health plans, publications in medical journals, and social media d) X-rays, retinal scans, fingerprints, and medical photos

e) Regulatory prerequisites

f) Data on biomarkers

g) Sales data from both in-patients and out-patients

h) Sales and marketing information

i) Customer testimonials.

SKILLS PERFORMED BY BIG DATA TECHNOLOGY

Data Quality Management; Understanding Information Operations; Decision Support Models Services include data warehousing and mining, predictive analytics, data visualization, data processing and governance, data science, cloud computing, and business model enhancement.

DATAMANAGEMENT TECHNOLOGIES / TOOLS USED TO LEVERAGE BIG DATA

To manage the vast amounts of complex data that are being generated at an increasing rate, the pharmaceutical and healthcare industries require tools, strategies, and software for data operations. Analytical software can analyse and carry out a wide range of tasks, including predictive modelling, deciphering intricate data patterns, enhancing hospital operations, supporting medical personnel, analysing data to comply with legal requirements, and streamlining pharmaceutical manufacturing procedures.

1.Apache Hadoop [4]

The Apache Software Foundation created the most widely used open-source technology in 2006.It is the only technology that enables the native form storage of all health care data. In fact, Hadoop stores enormous volumes of social media data. Terabyte and petabyte levels of data can be processed and analyzed using Hadoop. Hadoop handles the overheads associated with managing massive data collections, storage, and access while enabling highly fast parallel processing. Its capacity may be extended by adding more clusters. Big data processing requires Hadoop technology because of its high processing capacity at a faster pace, security against hardware failure, good data backup capabilities that store copies of data, dependability, affordability, and lack of need for pre-processing. However, Hadoop software operation is difficult because it calls for professionals and specialists. One of the largest problems facing health organizations is finding the right people to analyse large amounts of data. For most firms, it can be difficult and costly to find qualified individuals, such as trained data scientists (professionals with expertise in data mining, visualization, analysis, manipulation, and discovery). Hadoop technology is frequently utilized in genetics, cancer therapies, patient vitals monitoring, and the detection and avoidance of false claims when paying health insurance.

2. Apache Cassandra

It is a commercially available technology that is based on Dynamo and was developed by the Apache Software Foundation in 2008. It can store two million columns in a single row, enabling the use of a vast amount of data without the need for prior knowledge of data formatting. It features highly scalable columnar database storage in particular. Health care data analytics, including electronic health records and clinical trial reports, are flawlessly managed using Apache Cassandra.

3. Apache spark [5]

The most popular technology in the healthcare system is called Spark, and it was created by Apache Spark in 2014. Spark has a very quick rate of processing and integrating data from various sources. The primary usage of Apache Spark is predictive analysis of patient data, allowing for the early identification and notification of health hazards to patients. It supports efficient patient health monitoring, raising treatment standards, implementing the required precautions, and providing better care. Apache Spark is widely utilized in genomic research to expedite the process of discovering variations through sequencing. Additionally, it can recognize the genomic sequencing of hereditary characteristics linked to particular medical disorders.

www.ijprems.com
editor@ijprems.com

INTERNATIONAL JOURNAL OF PROGRESSIVE RESEARCH IN ENGINEERING MANAGEMENT AND SCIENCE (IJPREMS)
(Int Peer Reviewed Journal)
Vol. 04, Issue 09, September 2024, pp : 400-404

e-ISSN : 2583-1062
Impact Factor : 5.725

BIG DATA ANALYTICS TRANSFORMINGTHE PHARMACEUTICAL INDUSTRY [6]

Due to its increased volume, velocity, and veracity, big data necessitates data analytics for processing and integration, which increases demand, particularly in the pharmaceutical and other industries. Pharmaceutical businesses heavily depend on research and clinical data to determine treatment efficacy, test theories, and detect patterns of disease. Big data analytics and algorithms are used by pharmaceutical companies to manage clinical operations, medical research, hospitals, scientific research labs, R&D investigations, business model design, database design, inpatient and outpatient data management, improving sales, and developing marketing strategies.

1)Big data analytics in drug discovery [7]

Big data analytics is essential to the development of vaccines and new medications. The traditional approach, in which researchers would manually test samples of various chemicals in search of novel treatments, was superseded by predictive modelling and drug design utilizing data analytics. This process required a significant investment of time, money, and resources. Predictive modelling speeds up the drug development process by helping to understand toxic effects and drug-drug interactions.

2)Big data analytics in developing precision medicine

Personalized medicine can be designed and developed with the aid of big data gathered from multiple sources, including patient electronic health records (EHRs), demographic data, medical records, medical history, and genetic constitution. AI, software tools, and computer models can all be used to process all treatment patterns, results, reports, and records efficiently. This is also beneficial for post-market surveillance and clinical decision support systems (CDSS).

3)Big data analytics in enhancing the efficacy of clinical trials

For every pharmaceutical product or vaccination to ensure that the particular treatment is both safe and effective for human subjects, clinical trials are crucial. The identification of suitable trial patients is a crucial, costly, and time-consuming aspect of clinical studies. Consequently, to increase the success rate of clinical trials, pharmaceutical companies use big data to identify the right human subjects for inclusion in these trials by thoroughly evaluating the patient's genetic information, demographics, disease state, personality traits, behavioural patterns, and medical history. These aids pharmaceutical corporations in creating precision medicine, or personalized medicine, that is pertinent and unique to each patient.

4)Big data analytics in accelerating Research and Development [8],[9]

Large amounts of data are produced at several phases of the drug development process, from patient usage and medical records to drug discovery and design. Pharmaceutical companies employ data analytics to efficiently extract and use relevant information, supporting research and development.

5)Big data analytics in driving sales and marketing

Pharmaceutical businesses might explore new markets by interpreting real-world data from several sources, including social media, electronic health records, and medical records. This aids in the creation and use of fresh approaches to marketing and sales. More effective therapies are produced in the pharmaceutical business through big data analytics.

3. APPLICATIONS OF BIG DATA IN HEALTH CARE

1)Big data analytics in supporting preventive health [1],[10]

In the future, big data analytics can effectively be employed to stop diseases from developing and from spreading. Numerous individuals globally are afflicted with neurological disorders that are challenging to detect in their early stages. They collected information from 2,00,000 neurologically healthy patients over a 20-year period using big data analytics in order to find biomarkers for brain disease for the early detection and prevention of neurological illness. Reports on cognitive performance, eye retina scanning, EEG (Electroencephalogram), MRI (Magnetic Resonance Imaging), and neurological data are among the data gathered. Big data technology and tools can be utilized to integrate data, which will help with early disease and disorder prevention.

2)Big data in controlling adverse drug events [11]

Utilizing data from the Spontaneous Reporting System Database and other cutting-edge electronic-based technologies is part of the ADE's big data initiative. It is helpful in estimating the risk score, which aids in assessing the health of the patient and deciding on the degree of attention and priority assigned to each patient. This, in turn, aids in the early detection of unfavourable occurrences and the taking of necessary safeguards. Data mining and other analytical techniques can help make this more feasible. Large data sets are associated through data mining in the healthcare industry, which involves statistical analysis to compare patients' medical conditions and treatment regimens. This entails determining the most compassionate course of action following the identification of ADRs.

www.ijprems.com
editor@ijprems.com

INTERNATIONAL JOURNAL OF PROGRESSIVE RESEARCH IN ENGINEERING MANAGEMENT AND SCIENCE (IJPREMS)
(Int Peer Reviewed Journal)
Vol. 04, Issue 09, September 2024, pp : 400-404

e-ISSN : 2583-1062
Impact Factor : 5.725

3)Big data analytics in cancer treatments

Researchers can obtain valuable insights from patient databases sourced from various hospitals and universities, as well as biopsy reports, genomic data, and treatment records of cancer patients. These resources may aid in the identification of therapeutic approaches that yield precise outcomes.

As an illustration, a group at Massachusetts General Hospital (MGH) created technology that could forecast a patient's chance of getting breast cancer. These days, a large number of researchers from various institutions are utilizing AI and big data analytics and tools to enhance cancer treatments.

Telemedicine

Through telemedicine, big data is used to determine the optimal medication dose for patients in faraway locations. Telemedicine makes it possible to provide medication and patient monitoring in remote locations. It helps with public awareness, medical education, and illness surveillance.

To improve diagnosis and test outcomes, medical professionals rely on big data analytics to extract valuable information from vast amounts of medical data.

4. ADVANTAGES IN HEALTH CARE [12,13,14]

- Ensures public health by doing surveillance on public health, forecasting the onset of diseases, and acting quickly to prevent and control them.
- Helpful in enhancing the effectiveness of clinical diagnosis and research by utilizing a variety of statistical techniques and algorithms to create individualized treatment plans for patients.
- Assists regulators in fulfilling the demands for the creation of individualized and patient-specific therapies
- Patient follow-ups: real-time patient health monitoring through various biosensors, smartphone applications, smart pills, smart bottles, etc. allows for the analysis of medication and treatment efficacy. Big data analytics plays a key role in this process.
- Forecasting epidemics through monitoring illness outbreaks and slowing their rapid spread.
- Applied to diabetes, cancer, and uncommon diseases prediction and model building.
- Applied to forecast the recurrence, course, and therapy response of cancer.
- Selecting and carrying out the best course of action for patients.
- Monitors and enhances the health care system effectively.
- Encourages accountability in the health system.
- Enhances human life quality and guarantees patient safety.
- Creating new resources for consumers and doctors.
- Shortens the length and expense of treatment.
- Beneficial for enhancing creativity.
- Encourages clinical judgment.

5. REGULATIONS GOVERNING BIG DATA ANALYTICS

The pharmaceutical and healthcare industries are seeing a surge in the need for Big Data, which presents enormous hurdles for organizations in terms of data collection, integration, and storage.

The public may be impacted if these issues are not adequately resolved since they may have an impact on the consistency, quality, and accuracy of data. To evaluate big data's impact on healthcare, the European Commission formed the Big Data Task Force. However, in other nations, including the USA and India, the process of utilizing, developing, and regulating big data analytics has been slower. Although there are no clear legislation in the USA pertaining to big data, any firm, organization, or health care sector that wishes to use big data activities must abide by various laws, including privacy laws particular to their industry and location.

The US Federal Trade Commission (FTC) oversees data protection laws in the US, however most of these laws are state-level. There are currently no laws in India that specifically address data analytics. Patients' sensitive and private information is protected by laws like the Information Technology Act (IT Act) 2000 and the Information Technology Rules (IT Rules) 2011.

CHALLENGES

- Data Security risks.
- Privacy risks.
- Need for new technologies and need for human skills.
- Data ownership and Governance.

6. CONCLUSION

Major sources of health information include clinical decision support systems, wearable technology, and electronic health records. Regulatory bodies like the Food and Drug Administration (FDA) and European Medicine Agency (EMA) continue to approve these sources of information. These approvals serve as prime examples of how the big data analytics, artificial intelligence, and digital platforms of the pharmaceutical and healthcare sectors enable the collection of more accurate patient-related data. Large volumes of data created at an increasing rate from several sources are too much for traditional data operation systems to handle. Big data technology is hence beneficial. Big data analytics provides us with knowledge on creative and economical medical interventions. Big data analytics proved crucial during the Covid-19 outbreak, helping with everything from locating the infection source to managing and preventing the disease from spreading further. Big data analytics must be governed by rules because, despite its many benefits and broad use, the medical industry still faces issues with data security and privacy.

7. REFERENCES

[1] Shailaja P, Nikhila Pallavi D, Guna Sree B, Snehalatha G, Satya Ashok K. A Review on "Role of Big Data Analytics in Health Care Sector and Pharmaceutical Industry. International Journal for Research Trends and Innovation (IJRTI). 2022;7(7): e2456-3315.

[2] Wullianallur Raghupathi and Viju Raghupathi, "Big data analyticsin healthcare: promise and potential," February 2014.doi: 10.1186/2047-2501-2-3.

[3] "How Big Data in Health Care Influences Patient Outcomes"; home/blog/how big data in health care influences patient outcomes; July 7 2021.

[4] "Healthcare Big data projects, applications and examples," June 2022; ProjectPro

[5] J. Archenaa and E.A. Mary Anita, "Interactive Big data management in healthcare using Spark," February 2016.doi: 10.1007/978-3-319-30348-2_21.

[6] Tushar Sonal, "8 Ways pharmaceutical companies ensure success with Analytics," August 2019.

[7] "Big data Analytics in the pharmaceutical industry. How is big data analytics revolutionizing the pharma industry," May 2022; Intone.

[8] "How Big data means big opportunities for Pharma industry," January 2022, i2e consulting.

[9] Shikah J. Alsunaidi, Abdullah M. Almuhaideb, Nehad M. Ibrahim, Fatema S. Shaikh, Kawther S. Alquadaihi, Fahd A. Alhaidari, Irfan Ullah Khan, Nida Aslam, and Mohammad S. Alshaharani, "Applications of Big data Analytics to control COVID-19 pandemic," March 2021.doi:10.3390/s21072282.

[10] "How Big data analytics can support preventive health"; Features; Health IT Analytics; Xtelligent health care media; May 2022.

[11] C. Lee Ventola, "Big data and pharmacovigilance: Data Mining for adverse drug events and interactions," June 2018.

[12] "Big Data in Healthcare; Examples, Advantages and Disadvantages"; May 5,2022.

[13] William Goddard "Pros and Cons of Big Data"; IT Chronicles.

[14] Belle, Raghuram Thiagarajan, S.M. Reza Soroushmehr, Fatemeh Navidi, Daniel A. Beard and Kayvan Najarian, "Big data analytics in Healthcare," July 2015.doi:10.1155/2015/370194.