International Journal of Science and Research (IJSR)
ISSN: 2319-7064
SJIF (2022): 7.942

Big Data Analytics: Challenges and Applications in Healthcare

Fathma Mohammed[1], Madiha Naaz[2]

[1]B. Pharm, Student at Clinosol Research, Hyderabad, India

[2]Pharm. D, Student at Clinosol Research, Hyderabad, India

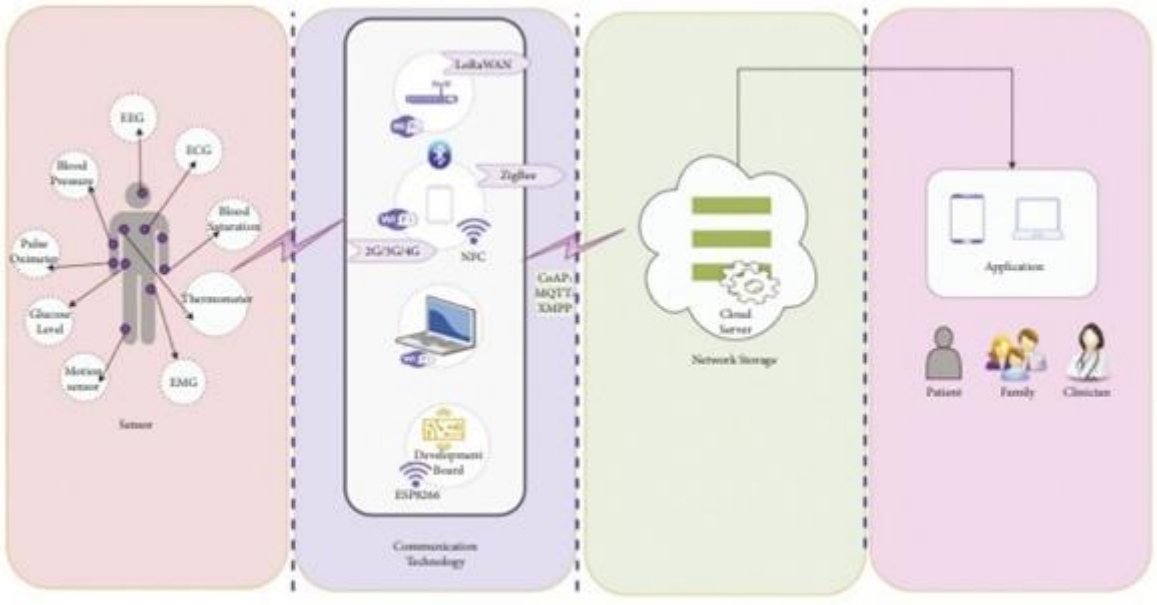
Abstract: ***Background: Although electronic health records, and other health – related technologies have proliferated in recent years, the collection of data in the healthcare industry has grown exponentially. A revolutionary tool for utilizing this plethora of health care data is big data analytics. This paper offers a thorough analysis of the difficulties and possibilities posed by the use of big data analytics in health care applications. Objectives: The primary objectives of this study are to critically evaluate the situation in the field of Big Data analytics in health care today and to highlight the major challenges that the researchers as well as practitioners must overcome. We also want to clarify the possibilities and possible advantages that big data analytics might bring to the health care sector. Methods: A comprehensive assessment of the research papers and articles that have been published in the last ten years was performed. Research terms, inclusion/exclusion criteria, and pertinent databases have to be found as part of the review technique. A thorough summary was then provided by categorizing the chosen research and combining their findings. Results: In accordance to the report, big data analytics in healthcare encounter a number of difficulties, including issues with connectivity, data privacy and security, and a shortage of qualified data scientists and health care workers. It also draws attention to a number of potential, such as the use of statistical analysis to prevent disease, enhanced clinical decision support systems, and improved patient care via customized medicine. The analysis also identifies new trends that have immense potential to transform health care delivery, such as the implementation of artificial intelligence and machine learning algorithms. Conclusion: The articles emphasizes the significance of tackling the difficulties associated with using Big Data analytics for healthcare while taking use of the numerous benefits it offers. Adopting advanced analytics techniques can result in improved patient outcomes, lower health care costs, and more effective health care delivery as the health care sector continues to develop. Multidisciplinary collaboration, strong data governance, and perpetual creativity are necessary to fully utilize the opportunities of Big Data analytics in health care applications. Researchers, medical practitioners, and administrators can use this study as a helpful resource to help them make well – informed decisions in the rapidly evolving field of health care analytics.***

Keywords: Big data analytics, Data - driven healthcare, opportunities and challenges in healthcare, Big data

1. Introduction

A new age of data - driven decision - making and patient - centered treatment has begun as a result of the integration of big data analytics into healthcare systems. This article gives a broad overview of the effects of big data analytics on people's health as shown through scholarly systematic reviews. We want to provide a thorough overview of the advancements made, difficulties encountered, and future opportunities in harnessing big data for healthcare purposes by combining findings from several systematic studies. (1)

Big Data, the general term for data sets of structured and unstructured data that are extremely large and complex so that the traditional software, algorithm, and data depositories are lacking to collect, process, break down, and store the data has come an intensely studied area in recent times. (2). The advanced logical technologies developed for Big Data have driven its operations in numerous areas similar as combating crime, business prosecution, finance, Global Positioning System (GPS), commerce, trip, civic informatics, meteorology, genomics, complex drugs simulations, biology, environmental exploration, and health care (3). Health care data are one of the driving forces of Big Data. With advanced data generation technology, there presents an exponential adding trend in the volume of data.



For illustration, Some kinds of medical equipment especially those which are continuously wearable record data and the high – speed data requires fast processing; in a specific data source, the value may be limited, but in the public sector, it may get to a maximized value through emulsion of electronic health records (EHRs) and electronic medical records (EMRs). CT examination abdomen in a subject, is a generous source of high measures data showing the abdomen with a high resolution data that it's too salutary in clinical settings and exploration for discovering abdominal features. Web/ mobile operations in health care have been expanded that enable cases to shoot their signs and symptoms to the provider; those operations contain abecedarian conditions, first aid, types of medicines, and also direct the case to the specialist. Health care system collects real - time biomedical signals (e. g., ECG, pulse oximeter, and blood pressure) in different places on mobiles, a health care operation is installed, and health data are accompanied for analysis and storehouse by apall calculating system in health care; big data can be represented with the backing of progressed information technology which observes information to make policy - making better; and a life map can be used to probe medical charges and population aging, which applies substantiation of policy - decision (4)

The article's main focus is on the results of using big data analysis to healthcare administration. It emphasizes the difficulties health care companies encounter when incorporating big data into their decision - making processes (5)

Major Sources and Types in Health Care:

In health care, big data encompasses various data types, including:

1) **Structured Data:** This includes data in a tabular format with well - defined fields, such as electronic health records (EHRs), which contain patient demographics, diagnoses, medications, and lab results (6).
2) **Unstructured Data:** This type includes textual information like doctors' notes, radiology reports, and patient – generated data like social media posts or wearable device data. (7)
3) **Semi - Structured Data:** Data that doesn't fit neatly into structured or unstructured categories. For instance, XML or JSON files used in healthcare interoperability standards like HL7 (Health Level Seven).
4) **Genomic Data**: Information about a patient's genetic makeup, including DNA sequences, gene expression data, and variations. The Human Genome Project is a notable reference in genomics. (8)
5) **Imaging Data:** Radiology images (X - rays, MRIs, CT scans, etc.) and pathology slides provide crucial visual information for diagnosis and treatment. (9)
6) **Clinical Trial Data:** Data generated during clinical trials, including patient outcomes, adverse events, and treatment efficacy. (10)
7) **Administrative and Billing Data:** Information related to insurance claims, billing, and administrative processes within health care organizations.
8) **IoT and Wearable Data:** Data from Internet of Things (IoT) devices and wearables like fitness trackers or smart watches, which monitor patient health continuously. (11)
9) **Patient – Generated Health Data (PGHD):** Data provided by patients themselves, such as symptom logs, wellness apps, and self – reported information. (12)
10) **Social Determinants of Health (SDOH):** Data related to a patient's environment, lifestyle, and socioeconomic factors that can impact health outcomes. (13)
11) **Behavioural Data:** Information about a patient's behaviours, such as diet, exercise, and substance use.
12) **Machine – generated Data:** Data generated by health care machines and equipment, including telemetry data from monitors and sensors.
13) **Research Data:** Data from medical research studies, including experiments, surveys, and clinical investigations. (14)

These data types are crucial for healthcare analytics, personalized medicine, predictive modeling, and improving patient outcomes. It's important to handle and analyze these data types effectively and securely to derive meaningful insights and advance healthcare practice

Big Data Analytics Applications In Healthcare

1) **Disease Prognosis and Identification**
Systematic evaluations demonstrate the tremendous potential of big data analytics for the prognosis and diagnosis of illnesses. such as diabetes, cardiovascular conditions, and cancer. These models aid in prompt interventions and better patient results.
2) **Individualized Medicine and Treatment**The use of big data analytics enhances the customized character of healthcare. Systematic evaluations have shown that doctors may individually adapt treatment approaches for patients by looking at genetic, lifestyle, and medical history information. Big data has made precision medicine possible, which has the potential to increase treatment effectiveness while reducing side effects.
3) **Management and Epidemiology of Public Health**
In dealing with public health emergencies and comprehending disease epidemiology, big data analytics are crucial. Researchers can track the spread of illnesses, forecast outbreaks, and influence public health plans by collecting data from sources including social media, electronic health records, and wearable technology. Systematic reviews emphasize the importance of these strategies in reducing the effects of threats to the global health (15)

Challenges of Big Data Analytics in healthcare:

1) Data Volume and Variety:
Big data is continuously changing, introducing challenges and problems caused by the rapid growth of healthcare data. Healthcare generates an large volume of data, including electronic health records (EHRs), medical imaging, genomics data, wearables, and more (16). The diversity and scale of this data pose significant challenges in terms of storage, integration, and analysis. Big data issues that generally happen in health care organizations are covered by four main categories

a) A huge amount of unstructured data are included in big clinical data like hand written data and natural language,
b) A reasonable degree of difficulty is brought by clinical big data's analysis
c) Integration
d) Storage (17)

2) Economic Challenges:
The medical field facilities of patients and health care providers such as doctors are dependent on subscriptions services. It inadvertently negatively impacts technology advancements in connection with this process (18)
The current challenges in data storage are mainly due to huge costs. Medical data costs arise mainly from 3 areas.

a) The huge amount of medical data is one of the reason of storage costs.
b) There are also costs associated with displacing them from one place to another as well as analyzing them.
c) Finally, the types of medical data type are unique, including numerical data that record various investigational tests, as well as various diagnostic images, records made by doctors and nurses, and even diagnostic speech, video, and other unstructured data. Unstructured data are more difficult to store, analyse, and manipulate when compared to structured data. They also, to a certain point, increase the cost of storage. (19)

3) Data Privacy and Security:
Healthcare data are highly sensitive and subject to strict privacy regulations, such as the Health Insurance Portability and Accountability Act (HIPAA) in the US. Ensuring privacy of patient data while enabling meaningful analysis remains a critical challenge. Health care data are more sensitive and centralized than other Big Data types. There are significant concerns regarding confidentiality. Big Data rises the risks to patient data for two reasons.
First is the risk of the data itself. The data can be copied and stored without any space and time constraints, and this feature is characterized by high risk and long – term risk under BigData conditions.
Second is the risk of BigData usage of technology. Under Big Data technology conditions, even if a Big Database uses unknown personal encrypted data, there is still a user identity that can be re – accessed by residual risk, and personal identities can be re - determined by data link technology because Big Data uses pseudonymized personal confidential data that have been hidden but retain a residual risk of re - identification (20)

4) Data Quality and Accuracy:
The accuracy and completeness of healthcare data are paramount. Inaccurate or incomplete data can lead to incorrect diagnoses and treatment decisions (21)

5) Limited data uniformity and Interoperability:
The current standards and technologies are not enough to meet the requirements of the integrative applications of health care Big Data. The difficulties are two types.

a) The data lack uniformity in standards, consistent description format, and presentation methods.
b) Different levels of structured, semi - structured, and unstructured data compilation are difficult.

Additionally, each database consists of different software and data formats, especially the latter makes data comparison, analysis, displacement, sharing, and other processes more difficult Healthcare systems often employ different standards and formats for data, making interoperability a major challenge. Effective data sharing and integration across systems are essential for comprehensive patient care (22)

6) Regulatory Compliance:
Complying with healthcare regulations, such as HIPAA and the General Data Protection Regulation (GDPR), while leveraging bigdata analytics, requires meticulous planning and robust security measures (23)

7) Ethical Concerns:
The use of patient data for research and analytics raises ethical questions regarding informed consent, data ownership, and data usage transparency (24)

8) Volume of data:
The enormous large volume of health care BigData in the terabyte (TB) level and even petabyte (PB) level is now beyond the storage capabilities of personal computers and network file sharing programs, thus establishing a new

mechanism to share is urgently needed (25)

9) Replace Medical Staff:
Big data presents a benefit to know about possible future health problems, but it has also a huge risk: the doctors might get replaced. Big Data is not so performing as it should, without a human touch, but it is feared that, if its use increases, patients would not go to the doctors, but use the technology and assume doctors authority. Big Data in the with respect of healthcare cannot be rejected because continuously more and more units and companies invest in this growing field. But, one should consider its drawbacks and realize a procedure safe for both doctors and patients. The question whether Big Data can direct medical staff replacement has appeared as a subject for discussion on major internet sites like Forbes, Fortune. (26)

2. Recommendations for Further Research

We suggest a few possibilities for future study in the area of big data analytics and health care, drawing on the findings of systematic reviews:

1) Data Sharing and Standardization
Future research should concentrate on creating standardized standards for secure data exchange across healthcare systems. As a result, collaborative research will be made easier, and the quality and variety of data that are accessible for analysis will improve.

2) Transparency and Precise AI
The creation of comprehensible AI models for healthcare should be a top priority for academics to foster adoption and foster confidence. This requires developing algorithms that offer transparent and understandable justifications for their choices, allowing doctors to make well – informed recommendations.

3) Longitudinal Research and Outcome Evaluation
Big data analytics can offer insights into the long - term consequences of interventions and therapies by including longitudinal data. For a more thorough knowledge of healthcare treatments, future research should prioritize the integration of real – world data with patient – reported outcomes.

4) Patient Empowerment and Ethical Frameworks
It is essential to create strong ethical frameworks that include concerns with consent, privacy, and data ownership. The future of healthcare will also depend on giving individuals the capacity to manage their own health data and take part in decision - making processes. (15)

3. Conclusion

The results of systematic evaluations demonstrate the irrefutable influence of big data analytics on people's health. An overview of the many applications of big data in healthcare is provided in this article while also acknowledging the challenges and ethical considerations that accompany its implementation. By considering the recommendations for future studies, researchers and practitioners can work collaboratively to harness the full potential of big data analytics, ultimately leading to improved healthcare outcomes and a more patient - centric approach to medical care. The various applications of big data in healthcare have been outlined in this article

References

[1] BorgesdoNascimentoIJ, MarcolinoMS, AbdulazeemHM, WeerasekaraI, Azzopardi - MuscatN, GonçalvesMA, Novillo – OrtizD Impact of Big Data Analytics on People's Health: Overview of Systematic Reviews and Recommendations for Future Studies JMedInternetRes2021; 23 (4): e27275doi: 10.2196/27275
[2] Asante - Korang & Jacobs, 2016; Kyoungyoung Jee & Gang Hoon Kim, 2013; Khoury & Ioannidis, 2014; Tan, Gao, & Koch, 2015
[3] Chen, M., Mao, S., & Liu, Y. (2014). Big data: A survey. Mobile Networks and Applications, 19 (2), 171–209
[4] Banan Jamil Awrahman, Chia Aziz Fatah, Mzhda Yasin Hamaamin, "A Review of the Role and Challenges of Big Datain Health care Informatics and Analytics", *Computational Intelligence and Neuroscience*, vol.2022, ArticleID5317760, 10pages, 2022. https: //doi. org/10.1155/2022/5317760
[5] Shabbir, M. Q., Gardezi, S. B. W. Application of big data analytics and organizational performance: the mediating role of knowledge management practices. *J Big Data* **7**, 47 (2020). https: //doi. org/10.1186/s40537 - 020 - 00317 - 6)
[6] Hillestad, R. et al. (2005). Can Electronic Medical Record Systems Transform Health Care? Potential Health Benefits, Savings, And Costs. Health Affairs, 24 (5), 1103 - 1117.
[7] Rajkomar, A. et al. (2018). Scalable and Accurate Deep Learning for Electronic Health Records. npjDigital Medicine, 1 (1), 18.
[8] Collins, F. S., & Varmus, H. (2015). A New Initiative on Precision Medicine. New England Journal of Medicine, 372 (9), 793 - 795.
[9] Gulshan, V. et al. (2016). Development and Validation of a Deep Learning Algorithm for Detection of Diabetic Retinopathy in Retinal Fundus Photographs. JAMA, 316 (22), 2402 - 2410.
[10] Ioannidis, J. P. A. (2005). Why Most Published Research Findings Are False. PLoS Medicine, 2 (8), e124.
[11] Steinhubl, S. R. et al. (2015). Effect of a Home - Based Wearable Continuous ECG Monitoring Patch on Detection of Undiagnosed Atrial Fibrillation: The mSToPS Randomized Clinical Trial. JAMA, 313 (24), 2471 - 2480.
[12] Woods, S. S. et al. (2016). Patient Experiences With Full Electronic Access to Health Records and Clinical Notes Through the My HealtheVet Personal Health Record Pilot: Qualitative Study. Journal of Medical Internet Research, 18 (6), e123.
[13] Adler, N. E., & Stewart, J. (2010). Health Disparities Across the Lifespan: Meaning, Methods, and Mechanisms. Annals of the New York Academy of Sciences, 1186 (1), 5 - 23.
[14] Ioannidis, J. P. A. (2005). Why Most Published

Research Findings Are False. PLoS Medicine, 2 (8), e124.

[15] Borges do Nascimento I, Marcolino M, Abdulazeem H, Weerasekara I, Azzopardi - Muscat N, Gonçalves M, Novillo - Ortiz D Impact of Big Data Analytics on People's Health: Overview of Systematic Reviews and Recommendations for Future Studies J Med Internet Res 2021; 23 (4): e27275URL: https: //www.jmir. org/2021/4/e27275 DOI: 10.2196/27275

[16] D. V. Dimitrov, "Medical internet of things and big data in healthcare, " *Healthcare Informatics Research*, vol.22, no.3, p.156, 2016.

[17] (B. K. Sarkar, "Big Data for Secure Healthcare System: A Conceptual Design, " *Complex Intell. Syst*, vol.3, no.2, pp.133–151, 2017.)

[18] L. Wang and C. A. Alexander, "Big data analytics in health care systems, " *International Journal of Mathematical, Engineering and Management Sciences*, vol.4, no.1, pp.17–26, 2019.

[19] Youssef, A. E. (2014). A framework for secure healthcare systems based on big data analytics in mobile cloud computing environments. The International Journal of Ambient Systems and Applications, 2 (2), 1 - 11

[20] (Liang Hong, Mengqi Luo, Ruixue Wang, Peixin Lu, Wei Lu, Long Lu, Big Data in Health Care: Applications and Challenges, Data and Information Management, Volume 2, Issue 3, 2018, Pages 175 - 197, ISSN 2543 - 9251, https: //doi. org/10.2478/dim - 2018 - 0014.

[21] Weiskopf, N. G., & Weng, C. (2013). Methods and dimensions of electronic health record data quality assessment: Enabling reuse for clinical research. Journal of the American Medical Informatics Association, 20 (1), 144 - 151

[22] Kruse, C. S., Stein, A., Thomas, H., & Kaur, H. (2016). The use of electronic health records to support population health: A systematic review of the literature. Journal of Medical Systems, 40 (11), 256.

[23] Wagner, A. K., Soumerai, S. B., Zhang, F., &Ross - Degnan, D. (2018). Segmented regression analysis of interrupted time series studies in medication use research. Journal of Clinical Pharmacy and Therapeutics, 27 (4), 299 - 309.

[24] Mittelstadt, B. D., &Floridi, L. (2016). The ethics of big data: Current and foreseeable issues in biomedical contexts. Science and Engineering Ethics, 22 (2), 303 - 341

[25] C. S. Kruse, R. Goswamy, Y. Raval, S. Marawi, Challenges and opportunities of big data in health care: A systematic review, Jmir Medical Informaticas, 4 (4) (2016), p. e38

[26] ALEXANDRU, Adriana & Radu, Irina - Miruna & BIZON, Madalina. (2018). Big Data in Healthcare - Opportunities and Challenges Informatica Economica.22.43 - 54.10.12948/issn14531305/22.2.2018.05.