InSight Bulletin: A Multidisciplinary Interlink International Research Journal
Peer Reviewed International, Open Access Journal.
ISSN : 3065-7857 / Website: https://ibrj.us / Volume-3, Issue-1 / January - 2026

Original Article

Analytics of Big Data in Healthcare Systems: A Review

Sheetal S. Sabale[1], Manasi C. Kolli[2]
[1,2]CES's Dr. Arvind B. Telang Senior College of Arts, Science and Commerce Pradhikaran, Nigdi Pune

Manuscript ID:
IBMIIRJ -2026-030124

Submitted: 08 Dec. 2025

Revised: 12 Dec. 2025

Accepted: 07 Jan. 2026

Published: 31 Jan. 2026

ISSN: 3065-7857

Volume-3

Issue-1

Pp. 124-128

January 2026

Correspondence Address:
Sheetal S. Sabale
CES's Dr. Arvind B. Telang Senior College of Arts,Science and Commerce, Pradhikaran, Nigdi Pune
Email: sheetal.sabale1978@gmail.com

OPEN ACCESS

Quick Response Code:

Web. https://ibrj.us

DOI: 10.5281/zenodo.18950886

DOI Link:
https://doi.org/10.5281/zenodo.18950886

Abstract

Due to increased digitalization and the widespread use of wearable technology, smart medical equipment, and electronic health records, big data analytics has emerged as a revolutionary force in contemporary healthcare systems. Large amounts of structured, semi-structured, and unstructured data with high volume, velocity, and variety are produced by the healthcare industry. Optimal resource allocation, increased operational efficiency, and better clinical decision support are all made possible by the effective analysis of this data. In order to extract valuable insights that enable precision medicine, preventive care, and improved patient outcomes, big data analytics combines sophisticated computer models, artificial intelligence, and real-time analytics. Drug development, illness prevention, diagnosis and treatment, hospital operations, and post-care management are just a few of the important healthcare sectors where applications are found. But major issues including data quality, privacy, interoperability, storage, and economical information retrieval still exist. Large-scale healthcare data management and processing heavily rely on technologies like Hadoop, HDFS, Map Reduce, and Apache frameworks. Complex system modelling and simulations are further supported by scientific programming and sophisticated analytics. Big data analytics has enormous potential to save healthcare expenses, cut down on errors, and enhance treatment quality despite obstacles. To fully exploit its benefits in healthcare systems, more study on data governance, infrastructure improvement, and integration techniques is necessary.

Keywords: *Big Data, Healthcare, Big data challenges.*

Introduction

The integration of data analytics into healthcare management has become a transformative force in modern healthcare systems. Rapid digitization and technological advancements have led to the generation of vast amounts of healthcare data, commonly referred to as big data, originating from electronic health records (EHRs), medical imaging, wearable devices, sensors, and patient-generated sources. When effectively analysed, this data provides valuable insights that support improved clinical decision-making, operational efficiency, and patient outcomes. Healthcare systems are undergoing a significant digital transformation driven by advancements in medical information systems, mobile health technologies, and the Internet of Things (IoT). At the same time, the industry is shifting from volume-based to value-based care, increasing the demand for efficient, data-driven practices among healthcare professionals.

However, the growing volume, variety, and complexity of healthcare data—often existing in structured, semi-structured, and unstructured formats—pose major challenges to traditional data management and analytics techniques. Existing healthcare data analytics tools are increasingly inadequate for managing large-scale, heterogeneous datasets while ensuring data quality, privacy, and security. Regulatory requirements, such as HIPAA, further complicate data access and sharing. This paper highlights the limitations of current healthcare data analytics approaches and emphasizes the need for advanced, scalable, and secure big data analytics techniques to improve data processing efficiency, support disease prediction and prevention, and enhance overall healthcare system performance. Data analytics in healthcare management has gained increasing attention as healthcare organizations face rapidly growing and complex data volumes. Researchers and practitioners recognize analytics as a critical tool for enhancing decision-making, improving clinical outcomes, and optimizing operational efficiency and resource allocation.

How to cite this article:
Sabale, S. S., & Kolli, M. C. (2026). Analytics of Big Data in Healthcare Systems: A Review. InSight Bulletin: A Multidisciplinary Interlink International Research Journal, 3(1), 124–128. https://doi.org/10.5281/zenodo.18950886

Literature Review

Historical Context and Evolution

Data analytics in healthcare evolved from early use of electronic health records and basic data systems focused on retrospective analysis of structured clinical data. With the emergence of big data technologies, healthcare organizations gained the ability to process vast amounts of structured and unstructured data, including medical imaging and genomic information. This shift enabled the transition from descriptive analytics to predictive and prescriptive analytics, supporting more accurate forecasting, informed decision-making, improved patient care, and greater healthcare system efficiency.

Data Collection

The gathering of information is the initial step in the data mining process. However, in order to acquire certain data as requested and use it properly, plans and concepts should be considered even before the data is collected additionally, [3] noted that inadequate data cleaning can lead to low-quality data collection, which causes many initiatives to fail and go over budget.

Key Concepts, Theories, and Frameworks

Several key frameworks underpin the use of big data in healthcare decision-making. Information Lifecycle Management (ILM) emphasizes effective data handling across its lifecycle to ensure accuracy, security, and relevance for analytics. Additionally, the Technology Acceptance Model (TAM) explains how perceived usefulness and ease of use influence healthcare professionals' adoption of analytics technologies, supporting successful implementation and utilization. Despite the extensive body of research on healthcare data analytics, critical challenges persist, including data quality degradation, privacy and security risks, lack of interoperability standards, and ethical concerns surrounding data-driven clinical and managerial decisions. The literature highlights the complex evolution of healthcare analytics, encompassing foundational frameworks and emerging technologies, while emphasizing the need for seamless integration into clinical workflows. These gaps justify further examination of core analytics methodologies, big data architectures, and their operational applications in healthcare decision-making systems.

The healthcare sector is widely considered as one of the most important industries in information technology [1]. More and more, information technology has been considered as a practice that facilitates healthcare performance through using data and information efficiently within the healthcare sectors. Therefore, [1] said that in order to understand the relation between information technologies and healthcare, we first need to understand what are the technologies used in healthcare. Over the past several years, information technology functions have evolved not only as a provider of technology services but also as a strategic provider that creates and integrates industry infrastructures that allow and guarantee service quality [2].

Healthcare Sectors & Big Data Analytics

1. **Big data storage and management**

Defining where and how data will be organized and stored when it is acquired is one of the most crucial aspects of handling and managing data. Since the data was organized and kept in data warehouses and relational databases after being extracted and loaded from many external sources, standard methods of storing and retrieving such data are no longer effective. But before this data can be used and function, it must be transformed and organized [4]. Addition, an agile database that can handle the data logically and through exchange of information in order to adapt to the rapid data evolution is absolutely necessary due to the large amount of data that has become available and the numerous data sources [5].

2. **Patients Role in Healthcare Analytics**

In order to achieve a high degree of efficiency and accuracy, this section looks at how individuals (and patients in particular) can enhance healthcare analytics by comprehending small and personal data and educating themselves on how to work with the healthcare data analytics [6]. When [7] coined the phrase "citizen science," he was referencing the situation in which unqualified, non-professional people are capable of conducting and supporting healthcare analytics systems. As a result, organizations will need to teach people how to self-monitor and follow up on their health information. Essentially, in order to do effective data analytics, we must first educate people on how to understand and acknowledge the significance of handling such data, such as how to handle breast cancer[8]. According to [9], medical personnel still need to evaluate and clarify this data in order to take it into consideration and record it, whatever how knowledgeable and talented our patients are to give us the information we need. Additionally, he stated that when the data has been verified and evaluated, we must determine how to modify an individual's behaviour, beginning with parents and guardians who are in charge of rearing their children.

3. **Connectivity between Healthcare Analytics System and Individuals (Medical Staff and Patients)**

Connectivity methodologies draw on past experiences using technology in daily life to produce ideas and thoughts from interconnected networks of minds [10]. Furthermore, [11] has added that healthcare analytics incorporates more than just technology and knowledge; it also takes into account people's attachment to and familiarity with medical care systems as well as their personal skills, such as the capacity to learn and adopt such systems in their lives. This is because different people, particularly older people, have different attitudes and reasons for rejecting such technologies.

4. **Healthcare Predictions and Decision Support System (DSS)**

In order to assess all possible health risks and project future healthcare procedures, predictive techniques employ the patient's medical history [12]. Based on [13], predictive techniques can be used through forecasting, saving time and money, by obtaining and examining historical patient information, diagnoses, and details from the databases.

5. **Role of Predictive Analytics in Medical Healthcare**

Since the findings of predictive systems enable treatments and actions to be implemented when all the hazards are detected in early stages, which helps to save costs, predictive analytics helps healthcare sectors reach a high level of effective general treatment and preventive care [14]. Additionally, [15] stated that in order to receive the required treatment at the appropriate time, people can work and help medical care by monitoring and updating their medical status. Since decision-making

systems in healthcare care sectors can be improved by concentrating on diagnosis of patients, behaviour, and prevention in order to reach a high level of care and improve healthcare economics, the technology era has substantially contributed value to the healthcare decision support system [17].

6. **Healthcare Prediction Examples Healthcare**

Analytics of Data Predictive analytics can be used in the healthcare industry in a multitude of ways, such as the successful delivery of medical care, which can be accomplished by employing a model that suggests algorithms to support medical treatment for interacting diseases, which may be observed in the behavior and interactions of patients. In order to deliver the right therapy for the right patient at the right time, another way to employ predictive analytics is to leverage software services and apps when combined with electronic healthcare records to assess diagnoses and verify results. Additionally, Imamura discovered that the association diagnostic approach may effectively retrieve desired information from large databases [18].

7. **Financial Factors in Healthcare Predictive Analytics**

The most important and evident outcome of employing such technology in the healthcare industry is its impact on expenses. One of the key factors that significantly impacts the cost of healthcare predictive analytics is information. Medical care systems are currently focusing on improving healthcare analytics performance and cutting costs by compressing unstructured clinical records and eliminating irregular data. Large amounts of data will therefore be easily and effectively managed and monitored [19]. According to[20], predictive analytics can help prevent and minimize inaccurate prediction costs and time because it lowers the cost of data sourcing by recognizing only the required and desired data, which is standardized, simplified, and accessible from historical clinical databases.

A study titled "Big Data and Analytics in Healthcare" by [20] and a few others was published in the Biomed International Journal. Big data is a collection of data items whose size, speed, kind, and complexity require the search, adoption, and creation of new hardware and software methods in order to process, analyze, and portray the data in a cost-effective way. Genomic data processing, physiological signal processing, and medical picture analysis were their main areas of interest. An outline of the analytical techniques used in medical image analysis to enhance the interpretability of the contents shown, various obstacles, and current strategies in the creation of monitoring systems that use both high fidelity Big data applications in genomics encompass a wide range of issues, including waveform data and discrete data from non-continuous sources. In the paper "Improving Healthcare Using Big Data Analytics," [21] talked about how Hadoop data processing is one of the most effective choices available right now and how it will provide you an advantage when analyzing the data. The purpose of this study was to present a workable computer methodology that uses big data and analysis to improve healthcare by enhancing accessibility, affordability, and healthcare science.

In their work "BIG DATA ANALYTICS Contribution to Healthcare," [22] and collaborators explore the value of applying big data analysis to patient care datasets for improved understanding of patient involvement, health management, and care coordination. It demonstrates their research in healthcare and how big data offers enormous applications and solutions to biomedical issues. It comes to the assumption that big data will be beneficial and bring about more improvements to the healthcare system than previously expected. Predictive analysis appears to have a greater impact on health research since scientific publications can go viral, help, and forecast a huge number of emergency clinical cases.

Applications of Data Analytics in Healthcare Decision-Making:

1. **Clinical Decision Support:**

By giving healthcare workers with real-time, evidence-based insights, data analytics improves clinical decision support systems. To find patterns and correlations, analytics algorithms examine large datasets, such as health information systems (EHRs), medical histories, and treatment outcomes. This helps healthcare professionals diagnose illnesses, forecast patient outcomes, and customize treatment regimens. Analytics-driven decision support solutions guarantee more informed and personalised therapeutic judgments by offering practical suggestions at the point of service.

2. **Allocation of Resource:**

Data analytics offers a strong framework for attaining efficiency, and managing resource allocation is essential to managing healthcare. Healthcare businesses can estimate demand and proactively allocate resources by analysing past patient data, admission rates, and resource consumption patterns. Hospitals can modify personnel numbers, bed capacities, and medical supply inventories by using predictive analytics to anticipate patient influxes. This planned approach guarantees that resources are distributed where they are most required, improves operational efficiency, and reduces bottlenecks.

3. **Efficiency in Operations in Healthcare businesses:**

Data analytics plays a major role in upgrading the general level of operational efficiency in healthcare businesses. Organizations can find inefficiencies and make focused improvements through looking at patient itineraries, workflow patterns, and resource usage. Analytics-driven process optimization improves patient throughput, shortens wait times, and simplifies administrative work. According to [23], predictive modeling also helps to ensure the smooth operation of healthcare facilities by anticipating equipment maintenance requirements and minimizing downtime.

4. **Delivery of Healthcare and Advancement of Quality:**

Data-driven insights support on going quality improvement in the provision of healthcare. Analytics makes it easier to measure against industry standards, monitor important performance metrics, and follow healthcare guidelines. Healthcare businesses can pinpoint areas for service delivery and patient satisfaction improvement by examining patient feedback. Further, analytics facilitate the identification of best practices, allow healthcare personnel to share knowledge, and promote a culture of on going learning and development [23].

Big Data Analytics Challenges in Health Care

1. **Privacy and Security:**

For people and businesses that possess information or data about individuals, goods, activities, etc., privacy and security are major concerns. Private and personal information may be included in medical records that healthcare providers collect from individuals. As a result, safeguarding patient data from hackers and harm requires extreme caution. Many of the technologies used in big data analytics and data processing are open source and lack certain security features.

2. **Storage and Processing Issues:**

The most obvious problem with big data is undoubtedly storing and processing the enormous volume of data. Due to the massive volume of data gathered and shared by social media, healthcare providers, corporate transactions, and stock markets, data grows dramatically every time a new storage technology is developed. Furthermore, this data is not only large in volume but also comprises a variety of data types that are produced on a regular basis. The largest obstacle to managing this massive amount of data is that current or conventional systems cannot handle and store data of this size and type due to the rate of data explosion. Thus, using cloud computing can help alleviate the storage issue.

3. **Data Ownership:**

One important and persistent issue with big data applications in healthcare and other fields is data ownership. Petabytes of medical records typically belong to the hospitals, government healthcare systems, or healthcare providers that developed them, but they do not own the information contained inside. However, patients feel that they are the owners of the data. If healthcare professionals do not obtain formal consent from patients before utilizing data for experiences or research purposes, the ownership disputes may be resolved through the legal system.

4. **Skills Requirement:**

A data analyst is a specialist whose job it is to gather, clean, visualize, and model or turn raw data into the information blocks that developers, marketers, and even healthcare practitioners need. The skills needed to work in the big data industry are one of the biggest obstacles to handling big data. Analytical skills are necessary to work with big data, according to a recent study that looked at these talents.

Conclusion

By supporting improved clinical judgments, effective resource management, and improved patient care, big data analytics is essential to the advancement of healthcare. It makes use of vast and varied healthcare data to offer customized and predicted treatment. Healthcare systems can completely benefit from big data analytics and achieve better healthcare outcomes with good implementation and on-going research, despite difficulties such data privacy, security, interoperability, and high costs.

Acknowledgment

The authors would like to express their sincere gratitude to CES's Dr. Arvind B. Telang Senior College of Arts, Science and Commerce, Pradhikaran, Nigdi, Pune, for providing the necessary facilities and academic environment to carry out this research work.

Financial support and sponsorship

Nil.

Conflicts of interest

The authors declare that there are no conflicts of interest regarding the publication of this paper.

References:

1. Wager, K.A., Lee, F.W. & Glaser, J.P. 2005, ManagingHealth Care Information Systems: A Practical Approach for Health Care Executives, Wiley, United states of America.
2. LeRouge, C., Mantzana, V. & Vance Wilson, E. 2007, 'Healthcare information systems research, revelations and visions', European Journal of Information Systems, vol. 16, pp. 669 – 671.
3. Chordas, L. (2001), 'Building a better warehouse', Best's Review, vol. 101, no. 11, p. 117.
4. Bakshi, K. 2014, 'Considerations for big data: architecture and approaches', Aerospace conference, IEEE, pp. 1-7.
5. Herodotou, H., Lim, H., Luo, G., Borisov, N., Dong, L., Cetin, F.B. & Babu, S. 2011, 'A self-tuning system for big data analytics', Innovation data system research, Starfish, pp. 261-272.
6. Luciano, SJ., Cumming, GP., Wilkinson, MD. & Kahana, E. 2013, 'The emergent discipline of health web science', J med internet res, vol. 15, no. 8, p. 166.
7. Swan, M. 2012, 'Crowdsourced health research studies: an important emerging complement to clinical trials in the public health research ecosystem'. J med internet res, vol. 14, no. 2, p. 46.
8. Hanoch, Y., Miron-Shatz, T. & Himmelstein, M. 2012, 'Genetic testing and risk interpretation: how do women understand lifetime risk results?', Judgment and Decision Making, vol. 5, no. 2, pp. 23-116.
9. Miron-Shatz, T., Bowen, B., Diefenbach, M., Goldacre, B., Muhlhauser, I. & Smith, RSW. 2011, 'Barriers to health information and building solutions', Better decisions: envisioning health care, vol. 6, no. 1, pp. 191-212.
10. McHorney, CA. (2009), 'The Adherence Estimator: a brief, proximal screener for patient propensity to adhereto prescription medications for chronic disease', Curr Med Res Opin, vol. 25, no, 1, pp. 215 38.
11. LexisNexis, 'Healthcare clinical analytics', United States, viewed 4 April 2015,
12. Loginov, MV., Marlow, E., & Potruch, V. 2012, 'Predictive modeling in healthcare costs using regression techniques'. ARCH 2013.1 proceedings, pp. 1-32.

13. Conley, E., Owens, D., Luzio, S., Subramanian, M., Ali, A., Hardisty, A., & Rana, O. (2008), 'Simultaneous trend analysis for evaluating outcomes in patient-centred health monitoring services', Health Care Management Science, vol. 11, no. 2, pp. 152-66.
14. Obenshain, M. (2004), 'Application of data mining techniques to healthcare data', Infection Control and Hospital Epidemiology, vol. 25, no. 8, pp. 690-695.
15. Cannon, M., & Tanner, M. (2007), Healthy competition: What's holding back healthcare and how to free it, Cato Institute, Washington, D.C.
16. Imamura, T., Matsumoto, S., Kanagawa, Y., Tajima, B., Matsuya, S., Furue, M., & Oyama, H. (2007), 'A technique for identifying three diagnostic findings using association analysis', Medical & Biological Engineering & Computing, vol. 45, no. 1, pp. 51-59.
17. Bertsimas, D; Bjarnadóttir, M; Kryder, J; Pandey, R; Vempala, S; Wang, G. (2008), 'Algorithmic prediction of health-care costs', Operations Research, vol. 56, no. 6, pp. 1382-1557.
18. Bradley, P., & Kaplan, J. (2010), 'Turning hospital data into dollars', HFM (Healthcare Financial Management), vol. 64, no. 2, pp. 64-68.
A. Belle, R. Thiagarajan, S. M. R. Soroushmehr, F. Navidi, D. A. Beard, and K. Najarian, "Big Data Analytics in Healthcare," Biomed Res. Int., vol. 2015, pp. 1–16, Nov. 2015, doi: 10.1155/2015/370194.
19. R. Sonnati, "Improving Healthcare Using Big Data Analytics," Improv. Healthc. Using Big Data Anal., vol. 6, no. 3, pp. 142–146, 2015.
20. M. Singh, N. Delhi, V. Bhatia, R. Bhatia, and D. Specialist, "Big data analytics," pp. 239–241, 2017.
21. Benzidia, S., Bentahar, O., Husson, J., & Makaoui, N. (2023). Big data analytics capability in healthcare operations and supply chain management: the role of green process innovation. Annals of Operations Research, 1-25.
22. Chernof, B., & Kaufman, R. L. (1997). Improving quality improvement: a data-driven assessment. Western journal of medicine, 166(2), 151.