International Journal of Technology 14(6) 1320-1333 (2023)
Received May 2023 / Revised September 2023 / Accepted September 2023

International Journal of Technology

http://ijtech.eng.ui.ac.id

A Review of Big Data Trends and Challenges in Healthcare

Laila Baloch[1], Sibghat Ullah Bazai[1], Shah Marjan[2], Farhan Aftab[1], Saad Aslam[3], Tse-Kian Neo[4*], Angela Amphawan[3]

[1]Department of Computer Engineering, Balochistan University of Information, Technology, Engineering and Management Sciences, Quetta 87650, Pakistan
[2]Department of Software Engineering, Balochistan University of Information, Technology, Engineering and Management Sciences, Quetta 87650, Pakistan
[3]Department of Computing and Information Systems, School of Engineering and Technology, Sunway University, Selangor 47500, Malaysia
[4]CAMELOT, Faculty of Creative Multimedia, Multimedia University, Cyberjaya 63100, Selangor, Malaysia

Abstract. The healthcare sector produces an enormous amount of complicated data from several sources, such as health monitoring systems, medical devices, and electronic health records. Big data analytics may improve healthcare by enabling more effective decision-making, improving patient outcomes, and reducing costs. To improve the operational efficiency of healthcare organizations, scientific studies must search for the standardization and integration of data analysis equipment and methods. This systematic literature review aims to provide current insights on the topic by analyzing a total of 60 relevant articles published between 2017 and 2023. The review explores the challenges and opportunities in using big data in healthcare, including data security, privacy, data quality, interoperability, and ethical considerations. The article also explores big data analytics' potential uses in healthcare, such as personalized treatment, disease prediction and prevention, and population health management. It provides significant insights for healthcare providers, researchers, and practitioners to make evidence-based decisions, as well as underlines the need for more research in this area to fully realize the promise of big data in healthcare.

Keywords: Artificial Intelligence; Big data analysis; Big data in healthcare; Healthcare

1. Introduction

The amount of data generated by several public and private sector industries have grown exponentially. As more individuals participate in the generation of Big Data (BD), it becomes essential to establish a clear definition of BD and comprehend its implications. According to (Dash *et al.*, 2019), BD refers to a massive amount of information generated, stored, and analyzed by diverse industries to enhance the services they offer. This data contains structured, semi-structured, and unstructured forms and originates from a multitude of sources, showing an increasing frequency and size. Due to the size and complexity of this BD, traditional data processing systems are inadequate for handling it, requiring the use of specialized tools and techniques to extract valuable insights from it. BD has become increasingly relevant in healthcare as it can improve patient outcomes, incre- ase efficiency, reduce costs and facilitate medical research. To achieve ethical data

*Corresponding author's email: tkneo@mmu.edu.my, Tel.: +60383125631
doi: 10.14716/ijtech.v14i6.6643

use, its implementation must be properly planned and executed (Adnan and Akbar, 2019a). BD, in this context, refers to enormous and complex datasets derived from several sources, including electronic health records, patient-generated data, and medical images (Batko and Ślęzak, 2022). The use of BD has the potential to completely transform the healthcare sector by advancing patient care, improving outcomes, and lowering costs while also advancing medical research (Jayasri and Aruna, 2022).

Big data is categorized using various "Vs," with three core ones being size, speed, and diversity. Gartner defines big data as information assets with substantial volume, velocity, and/or variety. These aspects demand cost-efficient and inventive approaches to information handling, enhancing understanding, decision-making, and process automation (Abdalla, 2022). The collection of "Vs" related to big data includes the following:

1. **Volume**: This is a significant challenge in Big Data Analytics, referring to the vast amount of data that needs to be processed.
2. **Velocity**: This relates to the speed at which new data is generated, posing the challenge of managing and analyzing data in real time.
3. **Variety**: This encompasses the heterogeneity of healthcare data, requiring the integration and analysis of diverse data types comprehensively.
4. **Variability**: This indicates the inconsistency of data, necessitating careful interpretation in different contexts.
5. **Veracity**: This concerns the quality and trustworthiness of the data.
6. **Visualization**: Representing data in a visually meaningful fashion poses a challenge in Big Data Analytics due to its complex and diverse nature.
7. **Value**: The goal of Big Data Analytics is to extract valuable insights and hidden knowledge from extensive datasets.
8. **Valence**: Refers to big data connectivity. Valence changes with time and volume; it can be predicted and modeled (Saggi and Jain, 2018).

Healthcare big data (BD) can be categorized into four primary sources. Firstly, medical data includes individual patient health information obtained from records, often collected from public health and medical records. Secondly, public health data encompasses publicly available health-related information and lifelong health records. Thirdly, medical images provide visual insights into the body's interior, while electrocardiogram recordings offer graphical representations of heart activity. Lastly, data from medical experiments and literature supports the evaluation of new treatments through research articles. Figure 1 provides a visual overview of these healthcare big data types (Roham, Gabrielyan, and Archer, 2021).

Big data analytics (BDA) uncovers patient data patterns for personalized treatment and population health monitoring, aiding disease detection. However, ethical, privacy and technical issues in BDA pose challenges. Addressing them is vital for healthcare improvement. Protecting patient data is essential. Despite challenges, BDA enhances patient care (Jayasri and Aruna, 2022; Imamalieva, 2022; Khanra *et al.*, 2020; Hassan *et al.*, 2019; Dash *et al.*, 2019).

Research concerning the use of BDA in healthcare often addresses common matters but lacks an in-depth analysis from high-quality studies. While many studies assess sources, technologies, benefits, and challenges, few evaluate the quality of the examined documents (Khanra *et al.*, 2020). Further investigation is required to identify common contexts where BDA finds application in healthcare. This study aims to provide comprehensive summaries of research topics, trends, challenges, and potential solutions related to the impact of big data on global healthcare.

The following section presents the detailed literature review, whereas section 3 discusses methodology and analysis procedure, sources of big data papers, phases of systematic literature, and protocol for systematic literature review. Section 4 presents the quality assessment. The next section answers our research questions and the latest trends/ innovations that can be used to address big data needs. Section 6 provides a comprehensive discussion of the findings, while the last section concludes this article.

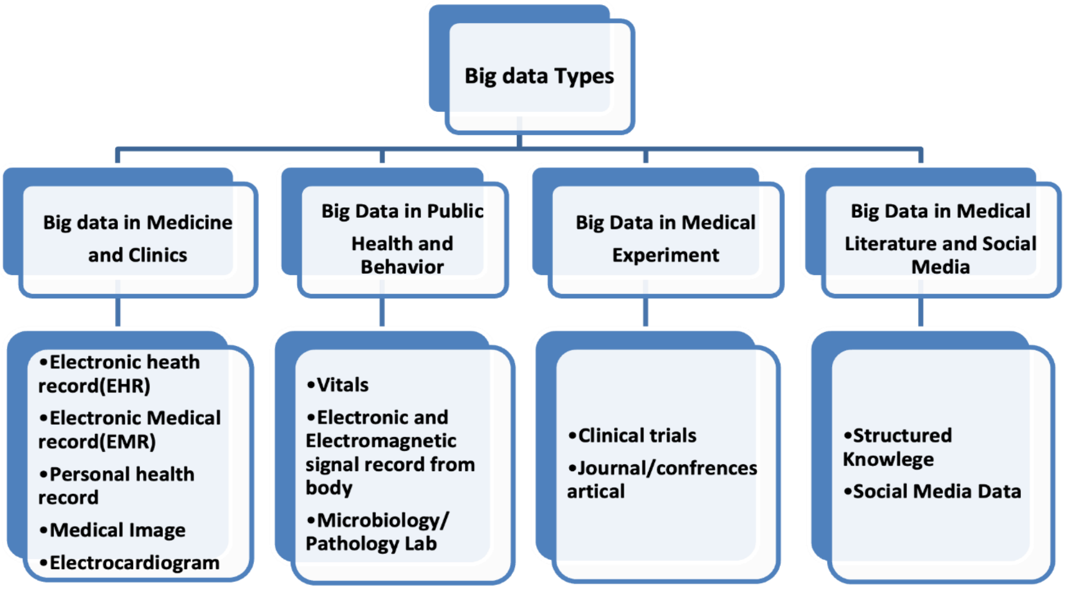

Figure 1 Illustrates the major types of BD in the healthcare industry

2. Literature Review

Big data analytics has numerous applications in healthcare with the expanding volume of big data in this context (Hassan *et al.*, 2019). BDA can be used to monitor the spread of diseases and predict outbreaks by analyzing large datasets from a variety of sources, including social media, public health monitoring systems, and electronic health records. BD can assist medical professionals in creating customized treatment strategies by analyzing data from genetic tests, medical history, lifestyle factors, and other sources to identify the most effective treatments for individual patients (Hassan *et al.*, 2022). Healthcare providers with real-time decision support by analyzing patient data and providing recommendations based on evidence-based medicine. To identify and stop healthcare fraud through the analysis of BD and identifying patterns that may indicate fraudulent activities.

2.1. Impact of big data on healthcare

The significance of BD in the healthcare system is immense, as it could change the way healthcare is managed, evaluated, and delivered. Here are some key reasons why big data is significant in healthcare (Cozzoli *et al.*, 2022).

Improving Patient Outcomes: BD aids healthcare practitioners in devising more efficient treatment plans and predicting potential health risks. Cloud-based remote patient monitoring models enable real-time health tracking, which is particularly beneficial for elderly patients living alone (Hassan *et al.*, 2019).

Improving Population Health: BD analysis of large datasets helps uncover trends and patterns that inform public health strategies and policies. It supports rapid decision-making during crises, such as tracking movements during the COVID-19 pandemic.

Optimizing Resource Allocation: BD optimizes resource usage, enhancing healthcare delivery by efficiently managing staff, facilities, and equipment. This leads to cost savings and improved effectiveness.

Personalizing Medicine: BD enables the creation of personalized treatment plans by analyzing patient data, including genetics and lifestyle factors. It guides individualized medical assessments and treatment strategies, improving outcomes and reducing risks (Hassan *et al.*, 2022).

Supporting Clinical Decision-Making: Real-time BD enhances clinical decisions, improving patient care. Integrating BD into healthcare is vital (Karatas *et al.*, 2022). BD, AI, and IoT aided COVID-19 tracking and health monitoring (Ahmed *et al.*, 2021). BD's impact spans personalized treatment, resource optimization, and real-time insights (Bag *et al.*, 2023).

2.2. Big Data management and analysis

Efficient big data analysis and decision-making require effective data management, particularly in critical fields like healthcare. In this context, three prominent platforms play significant roles: Hadoop, Apache Spark, and Apache Kafka. Hadoop, an open-source distributed application, excels at managing massive data across multiple machines, making it invaluable during data-intensive events such as pandemics (Harb *et al.*, 2020; Yuvaraj and SriPreethaa, 2019). Apache Spark is a versatile computing platform capable of handling vast datasets across various sectors, offering support for a wide range of data processing tasks. On the other hand, Apache Kafka is a distributed streaming platform that constructs real-time data pipelines and is crucial for real-time data processing and data exchange among different systems. These platforms collectively address the challenges of data management and computation in the era of big data (Harb *et al.*, 2020; Liang *et al.*, 2020; Dash *et al.*, 2019; Harerimana *et al.*, 2018).

3. Methodology

An SLR technique built on Kitchenham's concepts is applied to the proposed research (Kitchenham *et al.*, 2009). The process of the systematic literature review method is outlined in the following sections.

3.1. Research Questions

This review article addresses the following research questions:

RQ1. What is the significance of big data in healthcare?

RQ2. What are the research areas or trends that have emerged since 2017?

RQ3. What are the challenges and possible solutions to address these challenges?

3.2. Selection and analysis procedure

A snowballing approach is used as a starting point for conducting a literature study. Uses references and citations to identify additional sources, also known as backward and forward snowballing. Figure 2 illustrates the snowballing procedure.

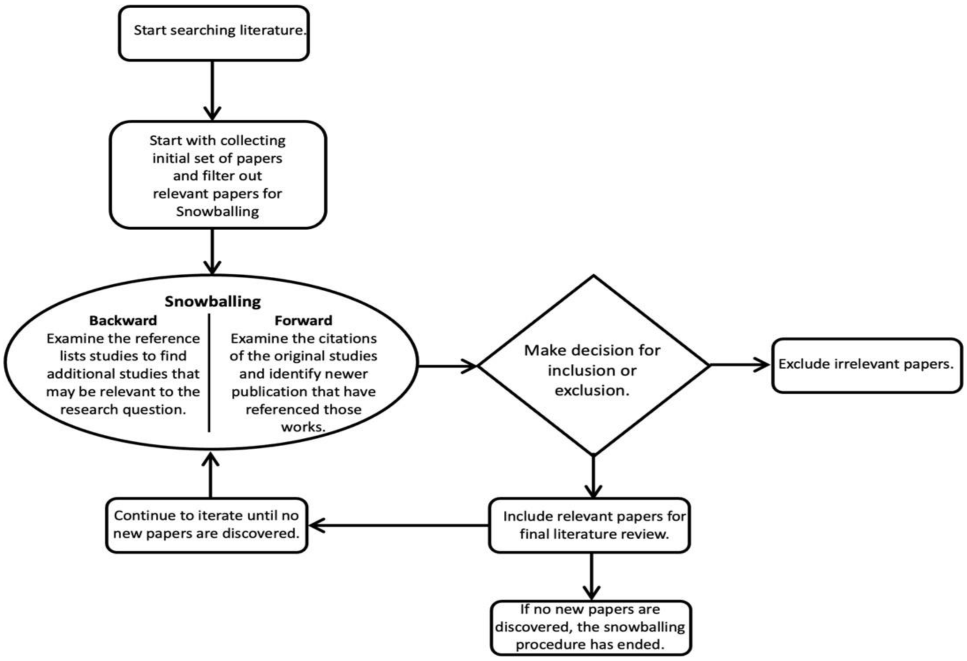

Figure 2 Snowballing procedure

3.3. Research Process

The analysis procedure is a manual search of conference proceedings and journal publications since 2017. 60 journals and conferences have been selected. Subsequently, we analyzed papers that addressed the literature survey questions, including the exclusion and inclusion of certain papers during this stage.

3.4. Sources of articles

Various academic databases, including Google Scholar, IEEE Xplore, ResearchGate, and Scopus, were utilized to search for relevant articles.

3.5. Inclusion criteria

We conducted a study focusing on healthcare-related articles pertaining to big data, data analytics, ML, and AI. Following the PRISMA guidelines (Rahmadian, Feitosa, and Zwitter, 2022), we filtered results from 2017 to 2023, specifically focusing on articles in the English language.

3.6. Exclusion criteria

- Studies that were not peer-reviewed or were not available in full text.
- Articles not written in the English language.

3.7. PRISMA flow diagram and Phases of systematic literature review

The PRISMA procedure was followed, as depicted in Figure 3, and the publications were filtered based on the following criteria:

1. During the 'Identification' stage, journal and conference papers relevant to the topic of big data were identified and subsequently included.
2. The screening step filtered out publications based on the titles and abstracts.
3. In the eligibility step, publications were excluded based on the whole text.

The result has a total of over 60 articles and proceedings available for analysis.

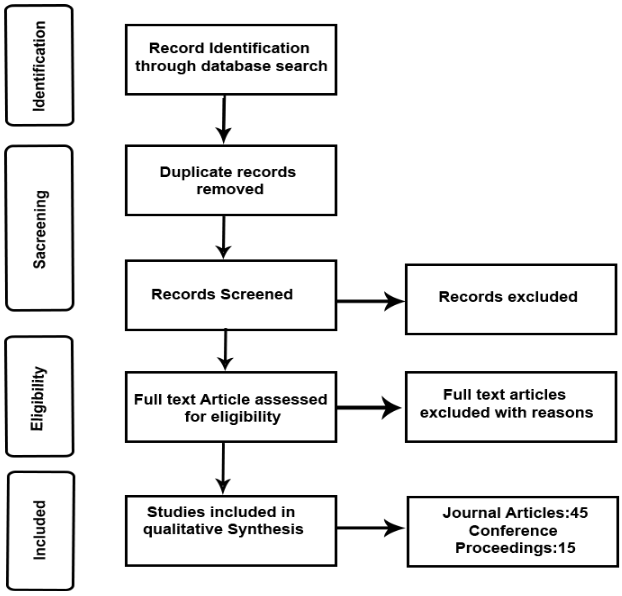

Figure 3 PRISMA flow diagram

3.8. Research Criteria

Analyze only scholarly works with healthcare and big data in title, abstract, or keywords. Focus on four aspects: data sources, approaches, purposes, and applications.

3.9. Protocol for systematic literature review

A standardized data extraction form collected author(s), publication year, title, thesis, method, results, and limitations. Data was synthesized and analyzed to identify recurring themes and patterns.

4. Quality Assessment

The studies' quality was assessed by considering factors like research question clarity, methodology appropriateness, and findings validity. This process aimed to identify unbiased studies to ensure reliable conclusions. The selected items were evaluated for transparency and impartiality, with a quality score calculated for each article using criteria from (Behera, Bala, and Dhir, 2019) and (Tandon *et al.*, 2020).

Table 1 Criteria for Quality Evaluation (QE)

QE#	Criteria
QE1	Analyses categorized as: quantitative = "+2", qualitative = "+1.5", no evidence = "+0".
QE2	Discussing advantages and challenges: Yes = "+2", Partial = "+1.5", No = "+0".
QE3	Results alignment with methodology: Yes = "+2", Partial = "+1.5", No = "+0". *Note: Partial explanation means inadequate method justification.*
QE4	Source reliability and peer review: "+2" if citations + H Index > 100. "+1.5" if citations + H Index between 50 and 99. "+1.0" if citations + H Index between 1 and 49. *"+0" if citations + H Index = 0.*
QE5	Comparison with prior methods: Yes (+1), No (+0).

Initially, 55 articles were found via keyword search, reduced to 45 after removing duplicates. Quality was assessed using a score (QS) from Table 2, with 12 studies scoring 9. Additional articles were found through citations. Those above 4.5 were relevant, while 9 and 8 scores were highly reliable for inclusion in the literature. (Behera, Bala, and Dhir, 2019).

Table 2 Computation of Quality Score (QS)

S.#	Reference	Total Citations	Quality Evaluation (QE)						Average citations
			QE1	QE2	QE3	QE4	QE5	QS	
1	(Jayasri and Aruna, 2022)	9	2	1.5	2	1	1	7.5	4.5
2	(Batko and Ślęzak, 2022)	32	2	2	2	1	0	7	32
3	(Liang *et al.*, 2020)	11	2	1.5	1.5	1	0	6	5.7
4	(Kim and Chung, 2019)	56	2	2	2	1.5	1	8.5	14
5	(Pham *et al.*, 2020)	226	2	1.5	2	2	0	7.5	75.3
6	(Deepa *et al.*, 2022)	237	2	2	2	2	1	9	237
7	(Wang *et al.*, 2023)	0	2	2	2	0	1	7	0
8	(Xing and Bei, 2020)	139	2	2	2	2	1	9	34.7
9	(Demirbaga and Aujla, 2022)	3	2	2	2	1	1	8	3
10	(Ghayvat *et al.*, 2022)	62	2	1.5	2	1.5	1	8	31
11	(Puthal, 2019)	28	2	1.5	2	1	1	7.5	5.6
12	(Zhou *et al.*, 2020)	17	2	2	1.5	1.5	1	8	5.6
13	(Li *et al.*, 2022)	23	2	2	2	1.5	1	8.5	23
14	(Nazir *et al.*, 2020)	66	2	2	1.5	1.5	0	7	22
15	(Kumar and Singh, 2019)	154	1.5	2	1.5	2	0	7	30
16	(Yan *et al.*, 2021)	6	2	1.5	2	1	1	7.5	1.5
17	(Sodagari, 2022)	0	2	1.5	1.5	0	1	6	0
18	(Imamalieva, 2022)	9	2	1.5	2	1	1	7.5	1.8
19	(Alexandru, Radu, and Bizon, 2018)	35	2	1.5	1.5	1	0	6	1.8
20	(Roham, Gabrielyan, and Archer, 2021)	4	2	2	1.5	1	1	7.5	0.75
21	(Adibuzzaman *et al.*, 2017)	52	2	1.5	2	1	1	7.5	10.4
22	(Hong *et al.*, 2018)	78	2	2	2	1.5	0	7.5	8.2
23	(Stylianou and Talias, 2017)	34	2	1.5	1.5	1	0	6	3.33
24	(Pastorino *et al.*, 2019)	153	2	2	2	2	1	9	22.25
25	(Lee and Yoon, 2017)	346	2	2	2	2	1	9	57.67
26	(Saggi and Jain, 2018)	242	2	2	2	2	1	8	48.4
27	(Hemingway *et al.*, 2018)	179	2	2	2	2	1	9	35.8
28	(Shilo, Rossman, and Segal, 2020)	220	2	2	2	2	1	9	73.33
29	(Zhang *et al.* 2017)	150	2	2	2	1.5	1	8.5	25
30	(Awrahman, Aziz-Fatah, and Hamaamin, 2022)	2	2	2	2	1	1	8	2
31	(Adnan *et al.*, 2020)	31	2	2	2	1	1	8	10.33
32	(Adnan and Akbar, 2019a)	136	2	2	2	2	1	9	34
33	(Adnan and Akbar, 2019b)	57	2	2	1.5	1.5	1	8	14.25
34	(Khanra *et al.*, 2020)	114	2	2	2	2	1	9	38
35	(Tandon *et al.*, 2020)	237	2	2	2	2	1	9	79
36	(Cozzoli *et al.*, 2022)	8	2	1.5	2	1	1	7.5	8
37	(Hassan *et al.*, 2022)	17	2	1.5	2	2	1	8.5	17
38	(Dash *et al.*, 2019)	771	2	2	2	2	1	9	192.75

39	(Al-Jaroodi, Mohamed, and Abukhousa, 2020)	65	2	1.5	1.5	1	1	7	21.67
40	(Khanna *et al.*, 2022)	2	2	2	1.5	1	1	7.5	2
41	(Yang *et al.*, 2020)	256	2	2	2	2	1	9	85.33
42	(Do-Nascimento *et al.*, 2021)	31	2	2	1.5	1	1	7.5	15.5
43	(Harerimana *et al.*, 2018)	75	2	2	2	1	1	8	15
44	(Ahmed *et al.*, 2021)	70	2	1.5	2	1	1	7.5	35
45	(Renugadevi, Saravanan, and Sudha, 2021)	10	2	2	2	1	1	8	5
46	(Syed *et al.*, 2019)	39	2	1.5	2	1	1	7.5	9.75
47	(Prosperi *et al.*, 2018)	138	2	2	2	2	1	9	27.6
48	(Boyapati *et al.*, 2020)	0	2	1.5	1.5	0	0	5	1.67
49	(Parimi and Chakraborty, 2020)	5	2	1.5	1.5	1	1	7	3
50	(Kuila *et al.*, 2019)	12	2	1.5	1.5	1	1	7	5

The number of yearly publications related to the topic has shown an upward trend from 2017 to 2020, as depicted in the chart (refer to Figure 4), particularly with increased academic interest in big data in 2020. Notably, a significant portion of the articles were published between 2018 and 2020. In both 2019 and 2022, there was an equal level of research interest.

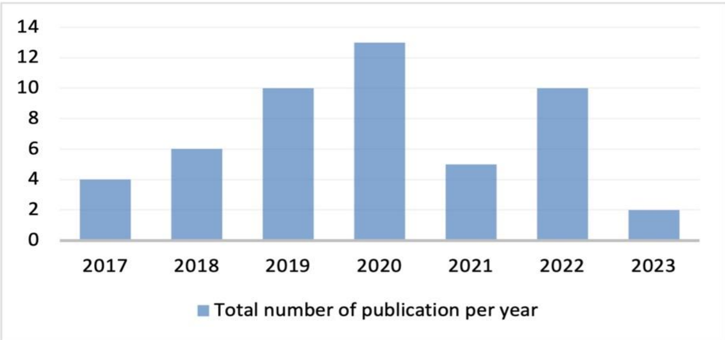

Figure 4 Yearly distribution of publications

5. Findings

5.1. Challenges and solutions

Big Data in healthcare holds immense potential but is confronted with issues such as data quality, privacy, governance, workforce shortages, and costs. Addressing these challenges can lead to improved patient outcomes, personalized healthcare, and more effective responses to public health crises, such as COVID-19.

5.1.1. Gap between costs and outcomes

A major challenge is the growing healthcare cost-outcome gap. Initiatives aim to address this by better managing research insights and evidence, reducing resource waste, and enhancing patient care through a "continuous learning healthcare system" (Lee and Yoon, 2017).

5.1.2 Unstructured Big Data in Healthcare

Unstructured big data from sources like social media and IoT devices lacks defined structure, making traditional processing challenging. Healthcare relies heavily on this data, but its complexity hinders accurate analysis. Information extraction (IE) tech is essential for distilling insights, and improved IE approaches are crucial, aided by advanced analytics like NLP, ML, and DL (Adnan *et al.*, 2020; Adnan *et al.*, 2019; Adnan and Akbar, 2019a).

5.2. Bigdata trends in healthcare

Predictions suggest that maintaining current healthcare delivery will become more challenging in the next two decades. The COVID-19 pandemic exposed issues with data analysis and prediction accuracy (Batko and Ślęzak, 2022). A text mining method is proposed to extract linked features from health data, improving value creation and data management (Kim and Chung, 2019).

5.2.1. Blockchain technology and AI

Blockchain technology in big data covers data management, secure sharing, provenance, and auditing, with future research focusing on efficient systems, AI/ML integration, and new use cases (Deepa *et al.*, 2022). Its applications include telemedicine, clinical trials, supply chain, and patient-centric healthcare, but challenges like scalability and governance persist (Wang *et al.*, 2023; Tandon *et al.*, 2020). The study highlights AI and big data's role in COVID-19, aiming to enhance predictions, diagnostics, tracking, drug discovery, and vaccine development (Pham *et al.*, 2020). Data mining, involving preparation, mining, and analysis, extract insights from extensive datasets (Yang *et al.*, 2020; Berawi, 2020).

5.2.2. Internet of Things

Wearable IoT generates extensive health data, requiring real-time security measures to prevent breaches. An unsynchronized sensor data analytics model is proposed (Jonny and Toshio, 2021; Puthal, 2019) using Storm and Spark to monitor health data (Renugadevi, Saravanan, and Sudha, 2021). IoT devices lack biological data but use big data analytics and ML to predict activities (Syed *et al.*, 2019), benefiting elder care, rehabilitation, and chronic disease management. Health 4.0, using techs like IoHT, medical CPS, cloud, BDA, ML, blockchain, and smart algorithms, faces privacy and security challenges (Naruetharadhol *et al.*, 2022; Al-Jaroodi, Mohamed, and Abukhousa, 2020). Table 3 summarizes the findings, highlighting privacy concerns and ongoing challenges.

Table 3 Literature Summaries are included in this review

Ref	Opportunities	Challenges	Findings
(Adibuzzaman *et al.*, 2017)	• Integrate smart infusion pumps. • Collaborate with stakeholders.	• Data quality, • privacy, and regulatory policies.	• Sharing data broadly. • Enhancing scientific research. • Improving drug interaction analysis.
(Alexandru, Radu, and Bizon, 2018)	• Improved quality of care. • Enhanced fraud detection. • Reduction in expenses. • Decreased waiting times for medical treatments.	• Privacy concerns alongside the replacement of medical professionals. • challenges in addressing privacy matters	• Identified the most important factors for using big data in healthcare.
(Hong *et al.*, 2018)	• Improved patient treatment. • Cost reduction.	• Data storage. • Data mining. • Data exchange. • Advancing medical research.	• Features of big data in healthcare. • Applications of big data in healthcare. • Analytic techniques used in healthcare big data.
(Stylianou and Talias, 2017)	• Cancer research. • Climate change.	• Privacy and ethical dilemmas. • Inadequate support for data transmission and visibility. • Market share loss.	• Three-dimensional model. • Obstacles and ethical issues. • Individuals impacted.

(Pastorino *et al.*, 2019)	• Enhancing early diagnosis. • Disease prevention. • Improving pharmacovigilance and patient safety. • Advancing precision medicine. • Reducing inefficiencies and improving cost-effectiveness.	• Ethical challenges: Privacy concerns, Personal autonomy implications, Impact on public expectations of transparency, trust, and fairness. • Legal challenges related to data access and analysis.	• Report on best practice initiatives in Europe. • Aim: Providing fresh data for clinical care and expediting public health surveillance. • Sectors: Cancer and public health.
(Hemingway *et al.*, 2018)	• Richer profiles of health and disease. • Accelerated disease understanding. • Discovery of new disease subtypes. • Holistic population and health system understanding.	• Data quality. • Knowledge of available data. • Legal and ethical framework. • Data sharing. • Disease definition standards. • Scalable science tools. • Multidisciplinary workforce.	• Identifying obstacles: data quality, available data knowledge. • Legal and ethical framework for data use. • Data sharing challenges. • Establishing and maintaining public trust. • Creating standards for disease definition.
(Awrahman, Aziz-Fatah and Hamaamin, 2022)	• Personalized care. • Strengthened patient-provider relationships. • Reduced hospital expenses.	• Real-time processing. • Data quality and security. • Privacy of healthcare data. • Heterogeneity of data. • Healthcare data standards.	• Significance of big data (BD) in healthcare. • Addressed challenges associated with BD utilization. • Focused on data aggregation challenges in healthcare.

6. Discussion

The literature review on big data (BD) in healthcare underscores the increasing importance of data-driven approaches for enhancing patient care, public health, and medical research. The utilization of big data analytics has led to significant advancements in patient outcomes and cost reduction. Technological advancements such as AI, IoT, ML, deep learning, and wearable sensors have amplified the application of big data analytics in the medical domain. Precision medicine could transform healthcare by customizing treatments according to an individual's genetic traits, unlocking significant potential for personalized care. Wearable technology aids real-time remote patient monitoring and early detection of potential health issues. Challenges associated with big data in healthcare encompass data quality, privacy concerns, interoperability, and ethical considerations. Blockchain integration is identified to enhance patient care, privacy, decision-making, and management efficiency. The quality and interoperability of healthcare data are pivotal, as inaccurate information could lead to incorrect diagnoses and treatment plans. Future research should concentrate on developing digital platforms and specialized applications based on big data analytics, including those dedicated to managing diagnostic images.

7. Conclusions

This study employs a systematic literature review to assess BD in healthcare and outline future directions. Articles are critically evaluated for their understandability, effectiveness, and scalability. Key areas of focus emerge from analyzing BD trends and challenges in healthcare. Future research directions include more efficient and scalable blockchain-based systems, integrating blockchain with AI and ML, and exploring novel blockchain applications. Current implementations of blockchain may not be able to handle the increasing volumes of data, but they do offer a secure and immutable ledger for healthcare records. Wearable devices are assisting in the monitoring of patients in a timely and continuous manner. Using blockchain technology, these devices are capable of securely transmitting real-time health data to healthcare providers, enabling early detection of health issues and more proactive treatment. Despite BD's benefits, concerns include data privacy, security, and quality, necessitating proper measures. Interoperability is vital for seamless data exchange among healthcare professionals. BD has the potential to improve patient outcomes and reduce costs. Future efforts will likely focus on machine learning, data mining, and blockchain to enhance decision-making, patient outcomes, and cost efficiency.

Acknowledgments

This work is being conducted at the Balochistan University of Information Technology, Engineering, and Management Sciences (BUITEMS) in Quetta, Pakistan. It is with great appreciation that we acknowledge the invaluable support received from Dr. Sibghat Ullah Bazai, Dr. Shah Marjan and all authers during the review process.

We would also like to thank the members of the Sunway Contactless Health Monitoring Project (STR-RCGS-E-CITIES[S]-004-2021) for their contribution and collaboration.

References

Abdalla, H.B., 2022. A Brief Survey on Big Data: Technologies, Terminologies and Data-Intensive Applications. *Journal of Big Data,* Volume 9(1), pp. 1–36

Adibuzzaman, M., DeLaurentis, P., Hill, J., Benneyworth, B.D., 2017. Big Data in Healthcare – The Promises, Challenges and Opportunities from A Research Perspective: A Case Study with a Model Database. In: *AMIA Annual Symposium Proceedings*, Volume 2017, p. 384

Adnan, K., Akbar, R., 2019a. An Analytical Study of Information Extraction from Unstructured and Multidimensional Big Data. *Journal of Big Data,* Volume 6(1), pp. 1–38

Adnan, K., Akbar, R., 2019b. Limitations of Information Extraction Methods and Techniques For Heterogeneous Unstructured Big Data. *International Journal of Engineering Business Management,* Volume 11, p. 1847979019890771

Adnan, K., Akbar, R., Khor, S.W., Ali, A.B.A., 2019. Role and Challenges of Unstructured Big Data in Healthcare. In: *Data Management, Analytics and Innovation*. Singapore: Springer

Adnan, K., Akbar, R., Khor, S.W., Ali, A.B.A., 2020. Role and Challenges of Unstructured Big Data in Healthcare. *Advances in Intelligent Systems and Computing,* Volume 1042, pp. 301–323

Ahmed, I., Ahmad, M., Jeon, G., Piccialli, F., 2021. A Framework for Pandemic Prediction Using Big Data Analytics. *Big Data Research,* Volume 25, p. 100190

Alexandru, A.G., Radu, I.M., Bizon, M.L., 2018. Big Data in Healthcare-Opportunities and Challenges. Informatica Economica, Volume 22(2), pp. 43–54

Al-Jaroodi, J., Mohamed, N., Abukhousa, E., 2020. Health 4.0: On the Way to Realizing the Healthcare of the Future. *IEEE Access,* Volume 8, pp. 211189–211210

Awrahman, B.J., Aziz-Fatah, C., Hamaamin, M.Y., 2022. A Review of the Role and Challenges of Big Data in Healthcare Informatics and Analytics. *Computational Intelligence and Neuroscience,* Volume 2022, p. 5317760

Bag, S., Dhamija, P., Singh, R.K., Rahman, M.S., Sreedharan, V.R., 2023. Big Data Analytics and Artificial Intelligence Technologies Based Collaborative Platform Empowering Absorptive Capacity in Health Care Supply Chain: An Empirical Study. *Journal of Business Research,* Volume 154, p. 113315

Batko, K., Ślęzak, A., 2022. The Use of Big Data Analytics in Healthcare. *Journal of Big Data,* Volume 9(1), p. 3

Behera, R.K., Bala, P.K., Dhir, A., 2019. The Emerging Role of Cognitive Computing In

Healthcare: A Systematic Literature Review. *International Journal of Medical Informatics,* Volume 129, pp. 154–166

Berawi, M.A., 2020. Managing Artificial Intelligence Technology for Added Value. *International Journal of Technology,* Volume 11(1), pp. 1–4

Boyapati, S., Swarna, S.R., Sharma, S., Agrawal, R., 2020. A Review of Bigdata and Machine Learning Techniques in Healthcare. In: *Proceedings of the 3rd International Conference on Intelligent Sustainable Systems, ICISS 2020*. Institute of Electrical and Electronics Engineers Inc, pp. 756–761

Cozzoli, N., Salvatore, F.P., Faccilongo, N., Milone, M., 2022. How Can Big Data Analytics Be Used for Healthcare Organization Management? Literary Framework and Future Research from A Systematic Review. *BMC Health Services Research*, Volume 22(1), pp. 1–14

Dash, S., Shakyawar, S.K., Sharma, M., Kaushik, S., 2019. Big Data in Healthcare: Management, Analysis and Future Prospects. *Journal of Big Data,* Volume 6(1), pp. 1–25

Deepa, N., Pham, Q.V., Nguyen, D.C., Bhattacharya, S., Prabadevi, B., Gadekallu, T.R., Maddikunta, P.K.R., Fang, F., Pathirana, P.N., 2022. A Survey on Blockchain for Big Data: Approaches, Opportunities, And Future Directions. *Future Generation Computer Systems,* Volume 131, pp. 209–226

Demirbaga, U., Aujla, G.S., 2022. MapChain: A Blockchain-Based Verifiable Healthcare Service Management in IoT-Based Big Data Ecosystem. *IEEE Transactions on Network and Service Management,* Volume 19(4), pp. 3896–3907

Do-Nascimento, I.J.B., Marcolino, M.S., Abdulazeem, H.M., Weerasekara, I., Azzopardi-Muscat, N., Goncalves, M.A., Novillo-Ortiz, D., 2021. Impact of Big Data Analytics on People's Health: Overview of Systematic Reviews and Recommendations for Future Studies. *Journal of Medical Internet Research,* Volume 23(4), p. e27275

Ghayvat, H., Pandya, S., Bhattacharya, P., Zuhair, M., Rashid, M., Hakak, S., Dev, K., 2022. CP-BDHCA: Blockchain-Based Confidentiality-Privacy Preserving Big Data Scheme for Healthcare Clouds and Applications. *IEEE Journal of Biomedical and Health Informatics,* Volume 26(5), pp. 1937–1948

Harb, H., Mroue, H., Mansour, A., Nasser, A., Cruz, E.M., 2020. A Hadoop-Based Platform for Patient Classification and Disease Diagnosis in Healthcare Applications. *Sensors,* Volume 20(7), p. 1931

Harerimana, G., Jang, B., Kim, J.W., Park, H.K., 2018. Health Big Data Analytics: A Technology Survey. *IEEE Access,* Volume 6, pp. 65661–65678

Hassan, M., Awan, F.M., Naz, A., Deandrés-Galiana, E.J., Alvarez, O., Cernea, A., Fernández-Brillet, L., Fernández-Martínez, J.L., Kloczkowski, A., 2022. Innovations in Genomics and Big Data Analytics for Personalized Medicine and Health Care: A Review. *International Journal of Molecular Sciences,* Volume 23(9), p. 4645

Hassan, M.K., El Desouky, A.I., Elghamrawy, S.M., Sarhan, A.M., 2019. Big Data Challenges and Opportunities in Healthcare Informatics and Smart Hospitals. In: *Security in Smart Cities: Models, Applications, and Challenges*, pp. 3–26

Hemingway, H., Asselbergs, F.W., Danesh, J., Dobson, R., Maniadakis, N., Maggioni, A., Van Thiel, G.J.M., Cronin, M., Brobert, G., Vardas, P., Anker, S.D., Grobbee, Di.E., Denaxas, S., 2018. Big Data from Electronic Health Records for Early and Late Translational Cardiovascular Research: Challenges and Potential. *European Heart Journal,* Volume 39(16), pp. 1481–1495

Hong, L., Luo, M., Wang, R., Lu, P., Lu, W., Lu, L., 2018. Big Data in Health Care: Applications and Challenges. *Data and Information Management,* Volume 2(3), pp. 175–197

Imamalieva, D., 2022. Recent Challenges of Big Data Application in Healthcare System. *In*:

International Conference on Multidimensional Research and Innovative Technological Analyses, Volume 2012, pp.121–124

Jayasri, N.P., Aruna, R., 2022. Big Data Analytics in Health Care by Data Mining and Classification Techniques. *ICT Express,* Volume 8(2), pp. 250–257

Jonny, K., Toshio, M., 2021. Building an Implementation Model of IoT and Big Data and Its Improvement. *International Journal of Technology*, Volume 12(5), pp. 1000–1008

Karatas, M., Eriskin, L., Deveci, M., Pamucar, D., Garg, H., 2022. Big Data for Healthcare Industry 4.0: Applications, Challenges and Future Perspectives. *Expert Systems with Applications,* Volume 200, p. 116912

Khanna, D., Jindal, N., Singh, H., Rana, P.S., 2022. Applications and Challenges in Healthcare Big Data: A Strategic Review. *Current Medical Imaging Reviews,* Volume 19(1), pp. 27–36

Khanra, S., Dhir, A., Islam, A.N., Mäntymäki, M., 2020. Big Data Analytics in Healthcare: A Systematic Literature Review. *Enterprise Information Systems,* Volume 14(1), pp. 878–912

Kim, J.C., Chung, K., 2019. Associative Feature Information Extraction Using Text Mining from Health Big Data. *Wireless Personal Communications,* Volume 105(2), pp. 691–707

Kitchenham, B., Brereton, O.P., Budgen, D., Turner, M., Bailey, J., Linkman, S., 2009. Systematic Literature Reviews in Software Engineering - A Systematic Literature Review. *Information and Software Technology,* Volume 51(1), pp. 7–15

Kuila, S., Dhanda, N., Joardar, S., Neogy, S., Kuila, J., 2019. A Generic Survey on Medical Big Data Analysis Using Internet of Things*. Advances in Intelligent Systems and Computing,* Volume 815, pp. 265–276

Kumar, S., Singh, M., 2019. Big Data Analytics for Healthcare Industry: Impact, Applications, and Tools. *Big Data Mining and Analytics,* Volume 2(1), pp. 48–57

Lee, C.H., Yoon, H.J., 2017. Medical Big Data: Promise and Challenges. *Kidney Research and Clinical Practice,* Volume 36(1), pp. 3–11

Li, C., Dong, M., Li, J., Xu, G., Chen, X.B., Liu, W., Ota, K., 2022. Efficient Medical Big Data Management with Keyword-Searchable Encryption in Healthchain. *IEEE Systems Journal,* Volume 16(4), pp. 5521–5532

Liang, Y., Quan, D., Wang, F., Jia, X., Li, M., Li, T., 2020. Financial Big Data Analysis and Early Warning Platform: A Case Study. *IEEE Access,* Volume 8, pp. 36515–36526

Naruetharadhol, P., Srisathan, W.A., Gebsombut, N., Wongthahan, P., Ketkaew, C., 2022. Industry 4.0 for Thai SMEs: Implementing Open Innovation as Innovation Capability Management. *International Journal of Technology,* Volume 13(1), pp. 48–57

Nazir, S., Khan, S., Khan, H.U., Ali, S., Garcia-Magarino, I., Atan, R.B., Nawaz, M., 2020. A Comprehensive Analysis of Healthcare Big Data Management, Analytics and Scientific Programming. *IEEE Access,* Volume 8, pp. 95714–95733

Parimi, S., Chakraborty, S., 2020. Application of Big Data & Iot on Personalized Healthcare Services Article in. *International Journal of Scientific & Technology Research,* Volume 9(3), pp. 1107–1111

Pastorino, R., De-Vito, C., Migliara, G., Glocker, K., Binenbaum, I., Ricciardi, W., Boccia, S., 2019. Benefits and Challenges of Big Data in Healthcare: An Overview of The European Initiatives. *European Journal of Public Health,* Volume 29, pp. 23–27

Pham, Q.V., Nguyen, D.C., Huynh-The, T., Hwang, W.J., Pathirana, P.N., 2020. Artificial Intelligence (AI) and Big Data for Coronavirus (COVID-19) Pandemic: A Survey on the State-of-the-Arts. *IEEE Access,* Volume 8, pp. 130820–130839

Prosperi, M., Min, J.S., Bian, J., Modave, F., 2018. Big Data Hurdles in Precision Medicine and Precision Public Health. *BMC Medical Informatics and Decision Making,* Volume 18(1),

pp. 1–15
Puthal, D., 2019. Lattice-Modeled Information Flow Control of Big Sensing Data Streams for Smart Health Application. *IEEE Internet of Things Journal,* Volume 6(2), pp. 1312–1320
Rahmadian, E., Feitosa, D., Zwitter, A., 2022. A Systematic Literature Review on The Use of Big Data for Sustainable Tourism. *Current Issues in Tourism,* Volume 25(11), pp. 1711–1730
Renugadevi, N., Saravanan, S., Sudha, C.M.N., 2021. Revolution of Smart Healthcare Materials in Big Data Analytics. In: *Materials Today: Proceedings,* Volume 81(2), pp. 87–1176
Roham, M., Gabrielyan, A.R., Archer, N., 2021. A Systematic Review of Knowledge Visualization Approaches Using Big Data Methodology for Clinical Decision Support. *Recent Advances in Digital System Diagnosis and Management of Healthcare*
Saggi, M.K., Jain, S., 2018. A Survey Towards an Integration of Big Data Analytics to Big Insights for Value-Creation. *Information Processing and Management,* Volume 54(5), pp. 758–790
Shilo, S., Rossman, H., Segal, E., 2020. Axes of a Revolution: Challenges and Promises of Big Data in Healthcare. *Nature Medicine*, Volume 26(1), pp. 29–38
Sodagari, S., 2022. Trends for Mobile IoT Crowdsourcing Privacy and Security in the Big Data Era. *IEEE Transactions on Technology and Society*, Volume 3(3), pp. 199–225
Stylianou, A., Talias, M.A., 2017. Big Data in Healthcare: A Discussion on The Big Challenges. *Health and Technology,* Volume 7(1), pp. 97–107
Syed, L., Jabeen, S., Manimala, S., Elsayed, H.A., 2019. Data Science Algorithms and Techniques for Smart Healthcare Using IoT and Big Data Analytics. *Smart Techniques for a Smarter Planet: Towards Smarter Algorithms,* pp.211–241
Tandon, A., Dhir, A., Islam, A.N., Mäntymäki, M., 2020. Blockchain In Healthcare: A Systematic Literature Review, Synthesizing Framework and Future Research Agenda. *Computers in Industry,* Volume 122, p. 103290
Wang, R., Xu, C., Dong, R., Luo, Z., Zheng, R., Zhang, X., 2023. A Secured Big-Data Sharing Platform for Materials Genome Engineering: State-Of-The-Art, Challenges and Architecture. *Future Generation Computer Systems,* Volume 142, pp. 59–74
Xing, W., Bei, Y., 2020. Medical Health Big Data Classification Based on KNN Classification Algorithm. *IEEE Access,* Volume 8, pp. 28808–28819
Yan, L., Huang, W., Wang, L., Feng, S., Peng, Y., Peng, J., 2021. Data-Enabled Digestive Medicine: A New Big Data Analytics Platform. *IEEE/ACM Transactions on Computational Biology and Bioinformatics,* Volume 18(3), pp. 922–931
Yang, J., Li, Y., Liu, Q., Li, L., Feng, A., Wang, T., Zheng, S., Xu, A., Lyu, J., 2020. Brief Introduction of Medical Database and Data Mining Technology in Big Data Era. *Journal of Evidence-Based Medicine,* Volume 13(1), pp. 57–69
Yuvaraj, N., SriPreethaa, K.R., 2019. Diabetes Prediction in Healthcare Systems Using Machine Learning Algorithms on Hadoop Cluster, *Cluster Computing,* Volume 22, pp. 1–9
Zhang, X., Pérez-Stable, E.J., Bourne, P. E., Peprah, E., Duru, O.K., Breen, N., Berrigan, D., Wood, F., Jackson, J.S., Wong, D.W.S., Denny, J., 2017. Big Data Science: Opportunities And Challenges To Address Minority Health And Health Disparities In The 21st Century. *Ethnicity and Disease*, Volume 27(2), pp. 95–106
Zhou, S., He, J., Yang, H., Chen, D., Zhang, R., 2020. Big Data-Driven Abnormal Behavior Detection in Healthcare Based on Association Rules. *IEEE Access,* Volume 8, pp. 129002–129011