ISSN 2320 - 2602

Volume 11 No. 12, December 2022

International Journal of Advances in Computer Science and Technology

Available Online at http://www.warse.org/IJACST/static/pdf/file/ijacst0111122022.pdf

https://doi.org/10.30534/ijacst/2022/0111122022

The Role of Knowledge Extraction from Big Data for Enhancing Healthcare Services

Adel A. Alzahrani, Hassan A. Alsereihy

King Abdulaziz Medical City- Jeddah, MNGHA, Saudi Arabia

King Abdulaziz University, Jeddah, Saudi Arabia

Received Date : November 10, 2022 Accepted Date : November 29, 2022 Published Date : December 07, 2022

ABSTRACT

All knowledge management methods and activities are meaningless unless knowledge is generated for effective various applications. Based on this fact, the aim of this article is to shed the light on the function of knowledge extraction from big data in improving healthcare services. The research derives various results after conducting this survey and examining several studies based on the descriptive analytical approach that is based on examining several scholarly papers from 2011 to 2022. The main achieved results revealed that most key benefits of using big data in healthcare include traceability, analytical capability, speeds to decisions capability, predictive analytics capability, interoperability capability, quicker clinical research progress, a more personalized approach to patient care, and earlier illness identification. On the other hand, high dimensionality, inconsistency, sparseness, distortion, and bias are some characteristics of healthcare big data that make their management and processing difficult. In addition, barriers to adopting big data technology include skepticism about its benefits, the necessity for a substantial initial investment, and difficulty reengineering healthcare systems and gaining approval for disruptive technologies. 2-1. Finally concluded, it was shown that knowledge extraction 2-2. from big data had altered healthcare practice and revolutionized the link between sickness and health.

Key words: Knowledge Extraction, Big Data, Challenges, Benefits, Healthcare, EHR, ECG, EEG, ICU

1- INTRODUCTION

While healthcare prices have been steadily growing [33], the quality of treatment offered to patients in Saudi Arabia has mostly stayed the same. Several studies have recently shown that researchers can cut death rates, healthcare expenses, and medical issues at different institutions by adopting contemporary healthcare technology[20][22][32][34]. Recent developments in information technology have greatly simplified the process of gathering numerous types of healthcare data [33]. Data plays an increasingly important role in modern medicine. According to a recent Big Data analysis, the potential value of healthcare data is close to $300 billion [21].

Recent advances in ubiquitous computing have already transformed how people access vital medical services. Particularly important to these computer systems is data analytics. When integrated into healthcare data, analytic solutions have tremendous potential to shift healthcare delivery from reactive to proactive. Analytics' influence in the healthcare sector will increase over the next few years. Through studying health records, we can deduce the underlying trends. In addition, it will aid doctors in creating a unique profile for each patient, from which they may properly predict the probability that the person will have a health problem shortly. Thus, the current study aims to shed the light on the role of knowledge extraction from big data to enhance healthcare services by presenting the challenges and advantages of using big data in the healthcare sector. The research aim try to solve the major problems given by:

Q1: What are the major challenges and benefits of big data analytics in the healthcare sector?

Q2: What is the role of big data analytic in extracting the knowledge of healthcare? **2- Literature review**

2. BIG DATA ANALYTICS BACKGROUND

In the 1990s, while contemplating visualization as a huge data challenge, [9] first coined the term "big data." [43]were the first authors to uncover key references to big data in the academic literature of computer science. Big data was first used in the context of statistics and econometrics by [12] in the year 2000. Douglas Laney of Gartner expanded upon the idea in 2001 but never published a study [18]. Big data refers to very massive and complicated datasets that strain the capabilities of conventional information processing methods [34]. Capturing, storing, sharing, and analyzing data and visualizing, updating, or querying information privately are the primary obstacles or challenges related to big data. As shown by [27], big data deal with massive amounts of information that cannot be stored in traditional databases, necessitating the selection of an alternate method for data extraction and processing.

Data analytics is a portmanteau of "data" and "analytics," where "data" refers to raw facts, statistics, and information and "analytics" indicates the use of multiple tools to evaluate this data, regardless of its size [34]. Although, all software designed to analyze data fall under the analytics umbrella[42]. "Big data analytics" refers to applying various methods to massive data. Since huge amounts of data come from various sources, big data analytics is used to discover previously unseen beneficial patterns and associations [34]. To put it another way, big data analytics is nothing more than examining, modeling, cleaning, and manipulating Big data to gather information and facilitate conclusion-making [34].

2.1 Structure of the healthcare big data analytics

Using the data life cycle paradigm, [40] created architecture for analyzing large amounts of data, beginning with its collection and continuing through its transformation and final use. The suggested best-practice big data analytics architecture is shown in Figure 1; it is composed of five main architectural layers given by: Data, aggregation of data, analytics, exploration of information, and management of data are all components of these layers. To better understand how to turn healthcare data from multiple sources into relevant clinical information, healthcare managers may take advantage of the big data analytics elements that are organized in logical levels to serve certain purposes.

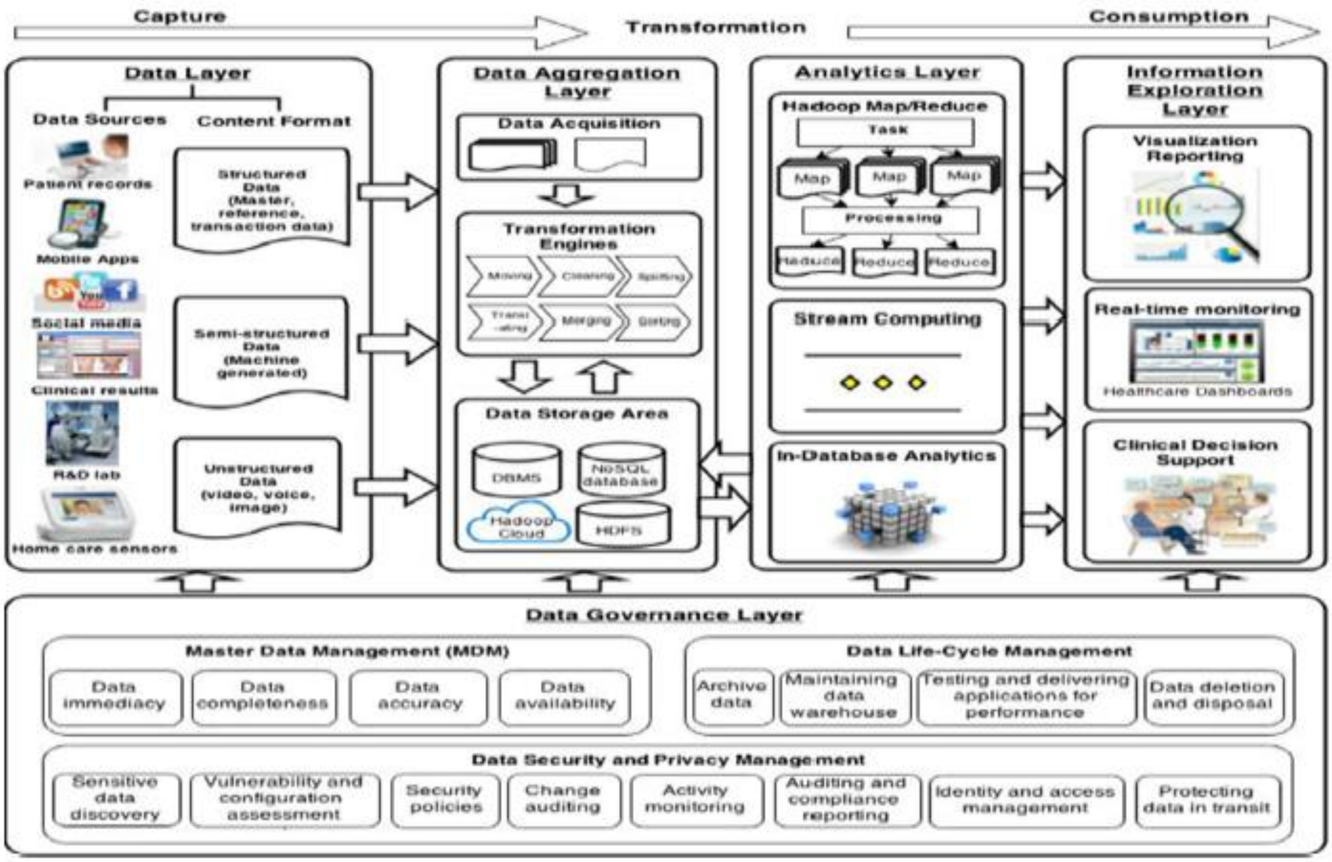

Figure 1: Structure of the healthcare big data analytics [40]

2.2 Data Sources and Basic Analytics in Healthcare

This section will focus on the effects of different data sources on analytical algorithms. Medical data mining sources are very diverse, necessitating a wide range of methods from various areas of data analytics. The various data sources of healthcare are given by:

- ***Electronic Health Records (EHR)***

A patient's medical history is recorded in EHRs or electronic health records. It includes everything from the patient's demographic to their medical history, test results, x-ray images, and billing information. In addition to a patient's medical or treatment history, many EHRs provide other, more holistic viewpoints on their care [33]. EHRs allow easy communication between healthcare professionals and institutions because of their useful features. EHRs are built for the modern healthcare environment, where authorized users constantly add and update data in real-time. The use of an EHR improves efficiency since staff may have instantaneous access to the most up-to-date patient information [15]. It helps with quality monitoring, patient outcomes reporting, and evidence-based decision support, all while creating a comprehensive record of a patient's clinical experience. EHRs simplify the archiving and retrieval of medical records. Improved care coordination, more precise diagnoses, better health outcomes, and a more positive patient experience are all benefits of this trend [11]

- ***Biomedical Image Analysis***

Medical imaging makes high-quality photographs of human anatomical features possible, which plays a crucial role in contemporary healthcare [33]. Magnetic resonance imaging, computed tomography, positron emission tomography, and ultrasound are the most common imaging modalities utilized to obtain a biological picture. The ability to examine human organs from the inside the body without causing harm to the patient has profound implications for medical practice. Without opening the patient physically, doctors may have a deeper understanding of what's happening within their body and why they're sick or injured [33]. Still, seeing these organs visually is just the beginning of the process. To get a deeper understanding of a medical condition, biomedical image analysis ultimately aims to collect quantitative information and draw conclusions from the pictures [33]. Research issues that occur while studying photos fall into many broad areas [33]. These include object identification, segmentation techniques, image restoration, and feature extraction. When these problems are fixed, healthcare data analytics will have the necessary inputs to provide useful analytical measures.

- ***Sensor Data Analysis***

Whether for diagnostic purposes in the present or historical research purposes in the past, sensor data [5] is omnipresent in the medical sphere. Many medical data-gathering tools, such as electrocardiograms (ECGs) and electroencephalograms (EEGs) are simply sensors that collect signals from different regions of the human body (1]. The most crucial use of real-time analysis is in the setting of intensive care units (ICUs) and the real-time remote surveillance of patients with particular medical problems [33]. The amount of information that has to be processed might be substantial in any of these scenarios. For instance, in an intensive care unit (ICU), the sensor may collect data from hundreds of sources, and alerts must be activated instantly [33]. Big data frameworks and specific hardware platforms are required for such applications. The current state of affairs and a more comprehensive study of trends and potential treatments over time are of significant interest in remote monitoring applications.

Despite the promising potential for increased healthcare efficiency brought about by the explosion of sensor data, the resulting data deluge poses a new difficulty. As a result, it is crucial to create cutting-edge data analytics tools capable of turning these massive datasets into actionable insights [33]. These analytic approaches will improve our ability to monitor patient's physiological signals and provide more situational awareness at the bedside [33]). They will also shed the light on the inefficiency of the health service that may be at the heart of rising costs.

- ***Biomedical Signal Analysis***

The biomedical signal analysis involves quantitatively evaluating biological signals with their roots in various physiologic procedures. Among the different types of such signals are the electrocardiogram (ECG), electromyogram (EMG), electrocardiogram (ECGs), electroencephalogram (EEG), electrogastrogram (EGGs), phonocardiogram (PCG), and others. It is crucial to analyze these signals to correctly identify pathological disorders and choose the best course of treatment [33].

Depending on the kind of treatment or severity of a medical condition, these signals might be either discrete or continuous. Low signal-to-noise ratio (SNR) and system interdependence make it difficult to analyze and understand physiological information [33].. Principal Component Analytical (PCAs), Singular Value Decomposition (SVDs), and Wavelet Transformations are just a few examples of the more advanced analytical approaches that have been extensively studied in the literature. These methods are also detailed in [3][4]. The techniques for analyzing time series are further upon in [24][29].

- ***Genomic Data Analysis***

Many illnesses have a hereditary component, but the exact relationship between certain genetic markers and illness is still unclear. Diabetes, for instance, is recognized to have a hereditary component; nevertheless, the whole collection of genetic markers predisposing a person to develop diabetes remains unclear [33]. The development of diverse gene treatments to heal these disorders will benefit greatly from a deeper understanding of the connections between specific genes, mutations, and diseases. Typically, data-driven studies are conducted with the primary goal of elucidating what kinds of health-related questions may be answered by analyzing genetic data in a computer [33]. Not only are those but there are also various obstacles to be overcome in the long and arduous process of putting genetic findings into tailored therapy. Genomic landscapes of illnesses like cancer, which include several interconnected factors, are notoriously complex and demonstrate a high order of variability across patients. If these problems can be resolved, it will be a great step toward making tailored medicine a reality [33]

Recent developments in biotechnologies have facilitated the quick production of vast quantities of medical and biological data, allowing for more in-depth studies in genomics. This has also opened up new possibilities and raised expectations for studying complex issues in life science on a genomic scale. For instance, the full genetic landscape of healthy people may now be studied for complicated disorders[38], thanks to advancements in genomic technology. Promising findings have been found in several of these research avenues, suggesting that they will lead to new understandings of the physiology of human illness and improved ability to predict individual response to therapy. Furthermore, genetic information is often represented as sequences or networks. Thus, proficiency in sequencing and network mining methods is essential for working in this area. Key research difficulties in medicine, such as identifying methods — and targeted therapies and predicting clinical outcomes, are being addressed by developing various solutions based on data analytics [33].

- ***Clinical Text Mining***

For the most part, patients' records are stored as text in the shape of clinical notes. Most healthcare information is based on these notes, often kept in an unstructured data format. Clinical data through dictation transcription, direct provider input, or voice recognition software are all included here. When it comes to untapped data resources, they may be the best. The human decoding of this free-text form on a wide variety of clinical information is, needless to say, too expensive and time intensive [33]. This is why it is restricted to main and secondary diagnoses and treatments for billing reasons. Clinical writing is often only provided in a free-form style, making it very difficult to interpret such notes mechanically [33]. Their lack of uniformity, variety in form and content, and contextual differences between patients and healthcare providers make this task challenging.

The fast and accurate automated encoding of clinical information relies heavily on natural language processing (NLP) techniques such as entity extraction [23]. Data preparation approaches are often more crucial than mining procedures in these situations. Short and telegraphic sentences, dictations, shorthand lexicons like acronyms and abbreviations, and often misspelled clinical terminology analyze clinical text using NLP approaches more difficult than processing other texts. The complexity of clinical text processing is increased because all these issues will indirectly affect the many typical NLP tasks like shallow or complete parsing, phrase segmentation, text classification, etc. [33].

- ***Mining Biomedical Literature***

Evidence from scientific publications is used in a wide variety of contexts. The latter is abundant and has increased dramatically over time. Biomedical applications that depend on data from scientific literature recognize the value of text-mining techniques for the long-term conservation, availability, and usefulness of digitally accessible resources. Innovative applications of new knowledge discovery techniques in the biomedical sector are possible with the help of text-mining methods and tools [23][46]. Supporting scholars in their quests for new insights, these technologies effectively search, extract, combine, analyze, and summarize textual data. The interdisciplinary character of biomedical text mining presents a significant hurdle. Biologists often use brand names to characterize chemical compounds, whereas scientists typically choose IUPAC-compliant names or unambiguous identifiers like International Chemical Identifiers [33]. Text mining techniques provide novel possibilities for efficiently populating, updating, and integrating such datasets, which would otherwise be prohibitively expensive to maintain. Biological research also benefits from text mining since it helps to generate ideas, save money on expert knowledge validation, and connect textual data to biomedical pathways [33]. The technique offers a standardized procedure for illuminating unrecognized connections and improving the structure of existing biological data.

- ***Social Media Analysis***

Rapid growth in social sources such as social networking websites, blogs/microblogs, platforms, query-answering services, and virtual communities has resulted in a plethora of data regarding public opinion on numerous healthcare issues. Data collected from social media platforms may be exploited for insights into health promotion and public health surveillance [33]. Public health experts may learn a great deal from the comments and suggestions of people on social media. Although most social media postings and messages are not very informative, when aggregated from millions of them, valuable insights may be gained [1][35]. Time lags associated with amassing such intricate data may be drastically cut in half with careful analysis of these massive hunks of information.

Capturing aggregate health patterns like infectious disease epidemics, reporting bad medication interactions, and enhancing interventional capacities are all areas where a prior study of social media insights for health has concentrated. Examining the evolution of social media material may provide light on illness outbreaks, making it a useful tool for early detection [33]. Online communities are a rich resource for information on a wide range of medical topics due to the high prevalence of common medical disorders among the general population. Due to the inherent unreliability of social media data, researchers must proceed with care when interpreting study findings [33].

3- RESEARCH METHODOLOGY

Descriptive studies are often classified as qualitative descriptions within qualitative research [17]. This research is often used in qualitative analyses of healthcare and nursing practices [30][31]. The who, what, and where of events or experiences, as well as gathering insights from informants about a poorly understood phenomenon, are all areas where qualitative description has been acknowledged as relevant and acceptable for research inquiries [17]. Qualitative description is appropriate when developing and refining surveys or treatments or when a straightforward explanation of phenomena is wanted [26][36] .

Sometimes, researchers in the health sciences conduct qualitative studies without formally documenting their approach. To distinguish itself from the more specific "methods," the article here uses the word "approach." This demonstrates a consistent epistemological perspective on the character of inquiry, the types of knowledge that may be found or created, and the appropriate methods for achieving these ends [14][16]. As a result, the current study depends on a qualitative descriptive approach to investigate various articles to determine the challenges and benefits related to using knowledge extracted from big data analytics in the healthcare sector.

4- ACHIEVED RESULTS & DISCUSSION

Healthcare information is being created exponentially and analyzed for insights every second. The healthcare sector accounts for around 30% of the world's data volume. Figure 2 illustrate that the compound annual growth rate of healthcare data will reach 36% by 2025. That's a rate that's 11% higher than the media and entertainment industry, 10% over the financial sector, and 6% above the industrial sector.

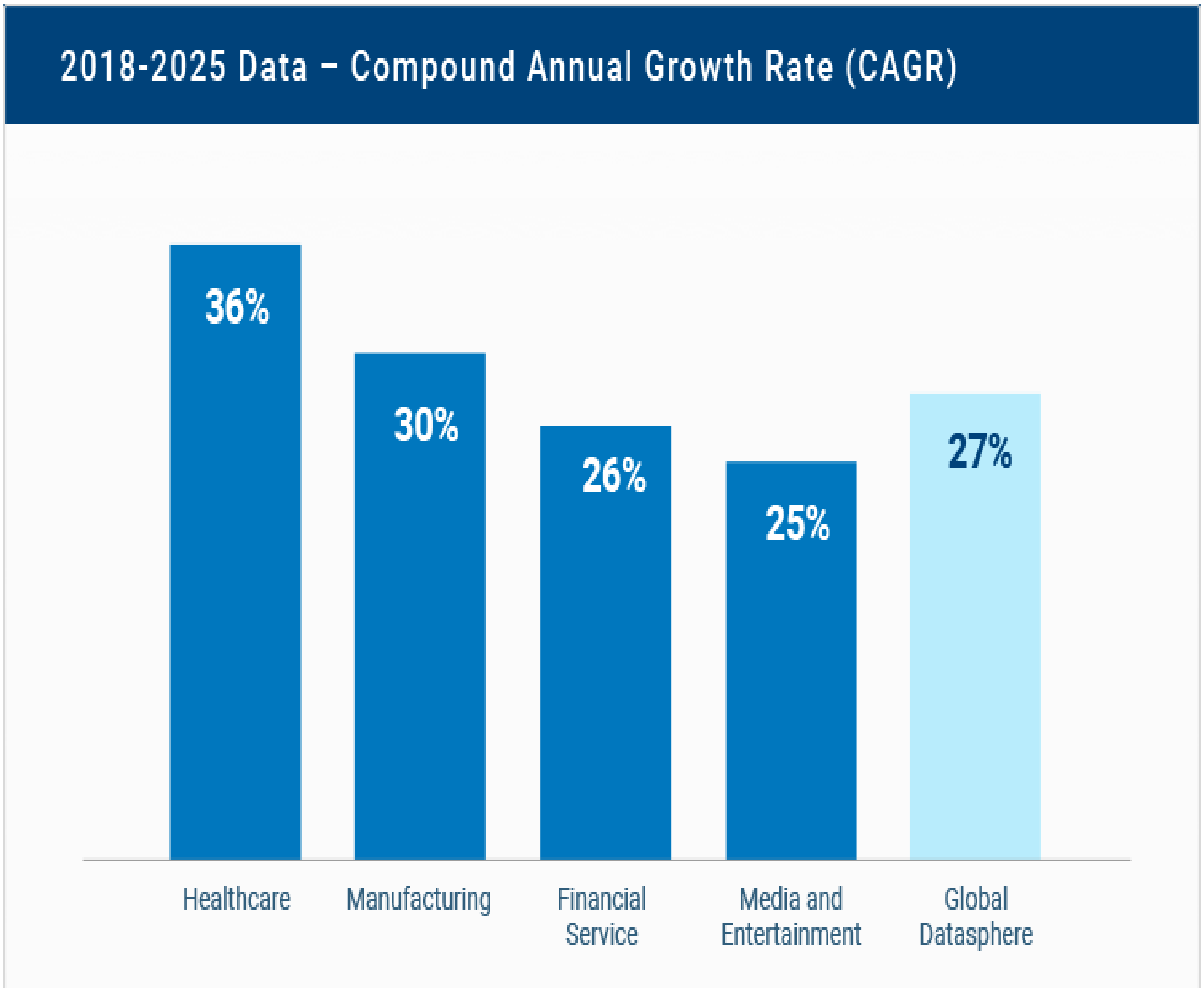

Figure 2: The compound annual growth rate of healthcare data [7]

In this section, the authors will describe the benefits and challenges that emerged in the investigated research. Figure 3 shows benefits and challenges of knowledge extraction from healthcare big data.

Figure 3: benefits and challenges of knowledge extraction from healthcare big data (Source: authors depend on previous studies).

4.1 Benefits of knowledge extraction from healthcare big data

The healthcare business has yet to realize the potential advantages of big data analytics. According to the descriptive study conducted by the authors, the most relevant benefits of extracting information from big data in the healthcare industry include:

- Possible advantages include early illness detection, which allows for more effective treatment, improved personal and community health management and the detection of healthcare fraud [32]
- Massive amounts of data can be used to estimate and predict things like the length of stay, the percentage of patients who opt for elective surgery, the percentage of patients who would not benefit from surgery, the likelihood of complications, the percentage of high-risk patients for medical complications, the likelihood of sepsis, MRSA, C. difficile, or any other care facility illness, the likelihood of illness progression, the percentage of patients at risk for progress in disease states, and the causes of illness [32]
- Traceability, analytical capability, speeds to decisions capability, predictive analytics capability, and interoperability capability [39].
- Possible advantages include quicker clinical research progress, a more personalized approach to patient care, and earlier illness identification [20]
- Benefits of IT infrastructure, operational benefits, managerial benefits, strategic benefits, and organizational benefits [39][40].
- Implementing big data techniques in healthcare settings can enhance the standard of care provided and save lives by allowing for the early detection of high-risk patients via basic analytics [22].
- Care quality and patient outcomes may be enhanced while keeping healthcare costs down. Extracting information from big data pave the way for a change from treating patients based on their own experiences to treating them based on scientific facts [22].
- The increased prices of health insurance benefits may be mitigated by implementing an electronic health record system that relies on big data to give meaningful data on the quality of treatment for the

participants of employee health insurance programs [10].

- Data obtained from IoT-reliant chips or sensors might shed light on important insights that could lead to better living, more efficient energy use, more convenient travel, and better health care [10].
- The healthcare industry stands to gain from applying the insights and actionable data made possible by the advent of big data, which will, in turn, facilitate the evolution of descriptive studies into prescriptive and predictive ones (Aceto et al., 2020).
- Finding connections and trends in this data aids medical professionals in providing high-quality treatment, saving lives and reducing healthcare costs [34..
- Big data is part of a healthcare organization's business intelligence approach to look at patient acceptance rates and employee productivity.
- Analyzing massive amounts of data, "big data analytics" may reveal previously unseen relationships and insights [34]
- Additionally, big data improves economic and administrative performance and readmission rates while decreasing pharmaceutical mistakes[34]

4.2 Challenges and limitation of knowledge extraction from healthcare big data

According to the authors' descriptive analysis, the following points represent the most important challenges of extracting knowledge from big data in the healthcare industry:

- Big data analytics platforms in healthcare should at least be able to perform the fundamental tasks involved in analyzing the data. Most of today's options are open source, but all include the usual caveats associated with that model [32]
- Unfortunately, healthcare data is seldom standardized, fragmented, and frequently produced in incompatible forms by antiquated IT systems [32]
- Data-related problems in the realm of large medical data include things like missing value, the constraint of diversity, and prejudice control [19].
- To achieve the potential of big medical data in the healthcare industry, several obstacles must be addressed, such as the need for more proof of the practical advantages of big data, technical concerns (such as legal and ethical issues), and clinical incorporation and usability issues [19].
- High dimensionality, inconsistency, sparseness, distortion, and bias are some characteristics of healthcare data that make their management and processing difficult [20]
- Despite the abundance of data, it is sometimes unclear which data should be used and why [37].
- Transitioning from paper-based records to dispersed data processing [28] [45] and a need for sufficient IT infrastructure (8] [13], Kung, & Byrd, 2018) also provide significant difficulties for the healthcare industry.
- Barriers to adopting big data technology include skepticism about its benefits [25][41], the necessity for a substantial initial investment[37] [44], and difficulty reengineering healthcare systems and gaining approval for disruptive technologies.
- From a technological perspective, problems arise when combining data with different levels of an organization and coming from different sources[6]
- Security and privacy of patient information are critical in the healthcare industry [22]
- Healthcare organizations face three main problems due to the rise in data volume, complexity, and velocity: storage, analysis, and financial stability[40]
- Implementing cutting-edge computing techniques, protocols, and hardware in the clinical environment is another significant challenge when dealing with healthcare data [10]
- Many developers of natural language processing tools find clinical documentation's specific nature and sophistication to be a formidable obstacle [10].
- Big data presents a significant challenge in the healthcare industry when received without an ideal data structure [10]
- Because of the severity of the consequences of a data breach, healthcare institutions must take precautions to protect their patients' personal information from hackers, phishing, and ransom ware [10]
- Extreme heterogeneity and analytical obscurity are the greatest challenges to extracting knowledge from large data in healthcare [2].
- Dealing with massive amounts of complex data presents unique difficulties in chemical analytics. Data sophistication, access to data, compliance with regulations, data security, optimal analytics methodologies, compatibility, administration, privacy, improvement, re-usability, open data, incomplete information, and data heterogeneity are all examples of such challenges [34].
- There are few limitations for this kind of research at the present time. First, although this research did conduct a descriptive assessment of a subset of relevant publications, it did not conduct a comprehensive analysis of the literature. So, future studies can use a systematic review to investigate a comprehensive analysis of the literature. Second, although this review provides an overview of the state of knowledge extraction from healthcare big data, it does not get into the specifics of how each research was implemented, how its conclusions were derived, and how challenges can be solved.

5- CONCLUSION AND FUTURE WORK

Improving healthcare delivery via knowledge extraction from big data is the next frontier. The information and digital revolution has provided the healthcare business with chances to reap the benefits of big data technologies while avoiding its challenges. The healthcare industry may benefit from knowledge extraction from big data in two main ways: by enhancing healthcare quality and results and by offering cost-effective treatment. Evidence-based medicine may replace experience-based medicine due to the prediction nature and pattern-recognition component of extracting information from big data. The study's descriptive method provides a solid foundation for incorporating big data analytics into future healthcare studies by elucidating the most salient benefits and difficulties of doing so. In addition, the research shows that using big data techniques would boost healthcare value by encouraging widespread use of insights after the extent of knowledge extraction from big data is defined, its characteristics and attributes are known, and challenges are adequately solved.

As a result, future studies can examine how challenges can be avoided. Third, the literature on knowledge extraction from big data and its advantages and problems in healthcare are fragmented, with information coming from a wide variety of different places. Finally, this review used a descriptive method; however the selection of papers on healthcare big data analytics was made on the basis of the authors' own preferences. So, further systematic reviews are needed.

REFERENCES

[1] Acampora, G., Cook, D. J., Rashidi, P., & Vasilakos, A. V. (2013). A Survey on Ambient Intelligence in Healthcare. *Proceedings of the IEEE*, *101*(12), 2470–2494. https://doi.org/10.1109/JPROC.2013.2262913

[2] Aceto, G., Persico, V., & Pescapé, A. (2020). Industry 4.0 and Health: Internet of Things, Big Data, and Cloud Computing for Healthcare 4.0. *Journal of Industrial Information Integration*, *18*, 100129. https://doi.org/https://doi.org/10.1016/j.jii.2020.100129

[3] Aggarwal, C. C. (2007). *Data streams: models and algorithms* (Vol. 31). Springer.

[4] Aggarwal, C. C. (2011). An Introduction to Social Network Data Analytics. In C. C. Aggarwal (Ed.), *Social Network Data Analytics* (pp. 1–15). Springer US. https://doi.org/10.1007/978-1-4419-8462-3_1

[5] Aggarwal, C. C. (2013). *Managing and mining sensor data*. Springer Science \& Business Media.

[6] Auffray, C., Balling, R., Barroso, I., Bencze, L., Benson, M., Bergeron, J., Bernal-Delgado, E., Blomberg, N., Bock, C., Conesa, A., Del Signore, S., Delogne, C., Devilee, P., Di Meglio, A., Eijkemans, M., Flicek, P., Graf, N., Grimm, V., Guchelaar, H.-J., … Zanetti, G. (2016). Making sense of big data in health research: Towards an EU action plan. *Genome Medicine*, *8*(1), 71. https://doi.org/10.1186/s13073-016-0323-y

[7] Capital Markets. (2018). The Convergence of Healthcare and Technology. *Capital Markets*. https://www.rbccm.com/en/gib/healthcare/episode/the_healthcare_data_explosion#content-panel

[8] Costa, F. F. (2014). Big data in biomedicine. *Drug Discovery Today*, *19*(4), 433–440. https://doi.org/https://doi.org/10.1016/j.drudis.2013.10.012

[9] Cox, M., & Ellsworth, D. (1997). Application-controlled demand paging for out-of-core visualization. *Proceedings. Visualization '97 (Cat. No. 97CB36155)*, 235–244. https://doi.org/10.1109/VISUAL.1997.663888

[10] Dash, S., Shakyawar, S. K., Sharma, M., & Kaushik, S. (2019). Big data in healthcare: management, analysis and future prospects. *Journal of Big Data*, *6*(1), 54. https://doi.org/10.1186/s40537-019-0217-0

[11] DesRoches, C. M., Campbell, E. G., Rao, S. R., Donelan, K., Ferris, T. G., Jha, A., Kaushal, R., Levy, D. E., Rosenbaum, S., Shields, A. E., & Blumenthal, D. (2008). Electronic Health Records in Ambulatory Care — A National Survey of Physicians. *New England Journal of Medicine*, *359*(1), 50–60. https://doi.org/10.1056/NEJMsa0802005

[12] Diebold, F. X. (2003). Big data dynamic factor models for macroeconomic measurement and forecasting. *Advances in Economics and Econometrics: Theory and Applications, Eighth World Congress of the Econometric Society,"(Edited by M. Dewatripont, LP Hansen and S. Turnovsky)*, 115–122.

[13] Fodeh, S., & Zeng, Q. (2016). Mining Big Data in biomedicine and health care. *Journal of Biomedical Informatics*, *63*(C), 400–403. https://doi.org/10.1016/j.jbi.2016.09.014

[14] Giorgi, A. (1970). *Psychology as a Human Science: a Phenomenologically Based Approach*. New York: Harper & Row.

[15] Hillestad, R., Bigelow, J., Bower, A., Girosi, F., Meili, R., Scoville, R., & Taylor, R. (2005). Can Electronic Medical Record Systems Transform Health Care? Potential Health Benefits, Savings, And Costs. *Health Affairs*, *24*(5), 1103–1117. https://doi.org/10.1377/hlthaff.24.5.1103

]16] Holloway, I. (2005). *EBOOK: Qualitative Research in Health Care*. McGraw-Hill Education (UK).

[17] Kim, H., Sefcik, J. S., & Bradway, C. (2017). Characteristics of Qualitative Descriptive Studies: A

Systematic Review. *Research in Nursing & Health*, *40*(1), 23–42. https://doi.org/https://doi.org/10.1002/nur.21768

[18] Laney, D., & others. (2001). 3D data management: Controlling data volume, velocity and variety. *META Group Research Note*, *6*(70), 1.

[19] Lee, C. H., & Yoon, H.-J. (2017). Medical big data: promise and challenges. *Kidney Research and Clinical Practice*, *36*(1), 3–11. https://doi.org/10.23876/j.krcp.2017.36.1.3

[20] Lee, C., Luo, Z., Ngiam, K. Y., Zhang, M., Zheng, K., Chen, G., Ooi, B. C., & Yip, W. L. J. (2017). Big Healthcare Data Analytics: Challenges and Applications. In S. U. Khan, A. Y. Zomaya, & A. Abbas (Eds.), *Handbook of Large-Scale Distributed Computing in Smart Healthcare* (pp. 11–41). Springer International Publishing. https://doi.org/10.1007/978-3-319-58280-1_2

[21] Manyika, J., Chui, M., Brown, B., Bughin, J., Dobbs, R., Roxburgh, C., Hung Byers, A., & others. (2011). *Big data: The next frontier for innovation, competition, and productivity*. McKinsey Global Institute.

[22] Mehta, N., & Pandit, A. (2018). Concurrence of big data analytics and healthcare: A systematic review. *International Journal of Medical Informatics*, *114*, 57–65. https://doi.org/https://doi.org/10.1016/j.ijmedinf.2018.03.013

[23] Meystre, S. M., Savova, G. K., Kipper-Schuler, K. C., & Hurdle, J. F. (2008). Extracting information from textual documents in the electronic health record: a review of recent research. *Yearbook of Medical Informatics*, *17*(01), 128–144.

[24] Mitsa, T. (2010). *Temporal data mining*. Chapman and Hall/CRC.

[25] Mohammed, E. A., Far, B. H., & Naugler, C. (2014). Applications of the MapReduce programming framework to clinical big data analysis: current landscape and future trends. *BioData Mining*, *7*(1), 22. https://doi.org/10.1186/1756-0381-7-22

[26] Neergaard, M. A., Olesen, F., Andersen, R. S., & Sondergaard, J. (2009). Qualitative description – the poor cousin of health research? *BMC Medical Research Methodology*, *9*(1), 52. https://doi.org/10.1186/1471-2288-9-52

[27] O'Reilly, T., Steele, J., Loukides, M., & Hill, C. (2012). Solving the Wanamaker Problem for Healthcare. *Big Data Now: 2012 Edition*, 83–96.

[28] Peek, N., Holmes, J. H., & Sun, J. (2014). Technical challenges for big data in biomedicine and health: data sources, infrastructure, and analytics. *Yearbook of Medical Informatics*, *23*(01), 42–47.

[29] Polasek, W. (2013). Time series analysis and its applications: With r examples. *International Statistical Review / Revue Internationale de Statistique*, *81*(2), 323–325. http://www.jstor.org/stable/43298885

[30] Polit, D., & Beck, C. (2009). International Differences in Nursing Research, 2005–2006. *Journal of Nursing Scholarship*, *41*(1), 44–53. https://doi.org/https://doi.org/10.1111/j.1547-5069.2009.01250.x

[31] Polit, D., & Beck, C. (2020). *Essentials of nursing research: Appraising evidence for nursing practice*. Lippincott Williams \& Wilkins.

[32] Raghupathi, W., & Raghupathi, V. (2014). Big data analytics in healthcare: promise and potential. *Health Information Science and Systems*, *2*(1), 3. https://doi.org/10.1186/2047-2501-2-3

[33] Reddy, C. K., & Aggarwal, C. C. (2015). An Introduction to Healthcare Data Analytics. In C. C. A. handan K. Reddy (Ed.), *Healthcare Data Analytics*. London, UK:

[34] Rehman, A., Naz, S., & Razzak, I. (2022). Leveraging big data analytics in healthcare enhancement: trends, challenges and opportunities. *Multimedia Systems*, *28*(4), 1339–1371. https://doi.org/10.1007/s00530-020-00736-8

[35] Sadilek, A., Kautz, H., & Silenzio, V. (2012). Modeling spread of disease from social interactions. *Sixth International AAAI Conference on Weblogs and Social Media*, 322–329.

[36] Sullivan-Bolyai, S., Bova, C., & Harper, D. (2005). Developing and refining interventions in persons with health disparities: The use of Qualitative Description. *Nursing Outlook*, *53*(3), 127–133. https://doi.org/https://doi.org/10.1016/j.outlook.2005.03.005

[37] Szlezák, N., Evers, M., Wang, J., & Pérez, L. (2014). The Role of Big Data and Advanced Analytics in Drug Discovery, Development, and Commercialization. *Clinical Pharmacology & Therapeutics*, *95*(5), 492–495. https://doi.org/https://doi.org/10.1038/clpt.2014.29

[38] Vogelstein, B., Papadopoulos, N., Velculescu, V. E., Zhou, S., Diaz, L. A., & Kinzler, K. W. (2013). Cancer Genome Landscapes. *Science*, *339*(6127), 1546–1558. https://doi.org/10.1126/science.1235122

[39] Wang, Y., & Hajli, N. (2017). Exploring the path to big data analytics success in healthcare. *Journal of Business Research*, *70*, 287–299. https://doi.org/https://doi.org/10.1016/j.jbusres.2016.08.002

[40] Wang, Y., Kung, L., & Byrd, T. A. (2018). Big data analytics: Understanding its capabilities and potential benefits for healthcare organizations. *Technological Forecasting and Social Change*, *126*, 3–13. https://doi.org/https://doi.org/10.1016/j.techfore.2015.12.019

[41] Wang, Y., Kung, L., Wang, W., & Cegielski, C.

G. (2018). An integrated big data analytics-enabled transformation model: Application to health care. *Information & Management*, *55*(1), 64–79. https://doi.org/https://doi.org/10.1016/j.im.2017.04.001
[42] Watson, H. J. (2014). Tutorial: Big data analytics: Concepts, technologies, and applications. *Communications of the Association for Information Systems*, *34*(1), 65.
[43] Weiss, S. M., & Indurkhya, N. (1998). *Predictive data mining: a practical guide*. Morgan Kaufmann.
[44] Wu, J., Li, H., Cheng, S., & Lin, Z. (2016). The promising future of healthcare services: When big data analytics meets wearable technology. *Information & Management*, *53*(8), 1020–1033. https://doi.org/https://doi.org/10.1016/j.im.2016.07.003
[45] Zhang, F., Cao, J., Khan, S. U., Li, K., & Hwang, K. (2015). A task-level adaptive MapReduce framework for real-time streaming data in healthcare applications. *Future Generation Computer Systems*, *43–44*, 149–160. https://doi.org/https://doi.org/10.1016/j.future.2014.06.009
[46] Zweigenbaum, P., Demner-Fushman, D., Yu, H., & Cohen, K. B. (2007). Frontiers of biomedical text mining: current progress. *Briefings in Bioinformatics*, *8*(5), 358–375. https://doi.org/10.1093/bib/bbm045