Migration Letters

Volume: 20, No: S1 (2023), pp. 2294-2304

ISSN: 1741-8984 (Print) ISSN: 1741-8992 (Online)

www.migrationletters.com

The Role And Challenges Of Big Data In Healthcare Informatics And Analytics: Review

Waad Saad Salem Allahyani[1], Rajwa Abdulaziz Adam Hawsawi[1], Reem Eqab Ayed Alharthi[1], Nisrin Mohammed Ismail Fallatah[1], Sinah Awad Safeer Aljauid[1], Abdulmajeed Safar Alosaimi[2], Faisal Kadimis Dahimesh Al-Dalbahi Al-Otaibi[3], Bander Muthyib Saleh Alotaibi[3], Majid Abdullah Musaed Alotaibi[4], Abdullah Saleh Musaed Al Otaibi[4], Fares Ibrahim Aida Alnafei[4], Ziad Muhammad Humayr Al Otaibi4

Abstract:

Background: *Data science is an interdisciplinary discipline that employs big data, machine learning algorithms, data mining techniques, and scientific methodologies to extract insights and information from massive amounts of structured and unstructured data. The healthcare industry constantly creates large, important databases on patient demographics, treatment plans, and results of medical exams, insurance coverage, and more. Moreover, healthcare has evolved with the development of technology to improve the quality of life and save lives. Today, big data is considered as one of the most essential and promising future technology areas and has been attracting the medical community's attention. As a result of big data, we can improve patient outcomes, personalize care, improve relationships between the patient and the provider, and decrease hospital costs. The effect of big data is very large since medical societies are known for their size, diversity of complexity, and a high degree of dynamism. Big data has been discussed from different viewpoints in recent years, protecting its involvement in many aspects, specifically those related to the healthcare system. Assembling health information, sharing data, and integrating health are essential in spreading health care. In addition, the security and privacy of data are critical since the data must be accessed from multiple locations within the distributed system.* ***This paper review aims*** *to recognize the role of big data in healthcare issues aggregating data and the challenges associated with big data in healthcare. The papers that have been selected for review are from last year's research.*

Introduction

Health informatics systems are indispensable implements for cultivating the eminence and effectiveness of healthcare. Health informatics systems are formed of databases and figures regarding all clinical facilities. The utilization of health informatics systems helps both healthcare providers and patients in the management of disorders and disease risks and in developing well-being [(1)]. However, there are a number of challenges to the application and practice of health informatics systems in the settings of primary health care [(2)]. Big data is well-known by its large qua[1]ntity and complexity and is produced from a wide range of sources, such as EHRs (electronic health records), medical imaging, genetic sequencing, and other sources. The rising popularity of DHTs (digital health technologies) and the need for more evidence-based healthcare decisions have both contributed to the rapid expansion of large amounts of data in the digital health industry in recent years [(3)].

Appropriate medical management for specific diseases will improve patient outcomes and reduction life-threatening conditions. It also decreases the side effects of drugs that influence their lives and medical waste products. Finding new drugs and equipment leads to further accurateness in the healthcare system [(4)]. Particular kinds of medical

[1]Health informatics specialist, Maternity and Children's Hospital in Makkah, Saudi Arabia.
[2]Health informatics technician, Dawadmi General Hospital - Third Health Cluster in Riyadh, Saudi Arabia.
[3]Health informatics technician, Rafa'a Al-Jamsh Hospital, Saudi Arabia.
[4]Health informatics specialist,Taif - King Abdulaziz Specialist Hospital, Saudi Arabia.

equipment especially those which are continuously wearable record data and the high-velocity data needs fast processing; in a specific data source, the worth may be limited, but in the public sector, it may get to a maximized value through fusion of electronic health records (EHRs) and electronic medical records (EMRs) [5]. CT scan for visualizing a patient's body, for instance, a patient's abdomen, is a plentiful source of high measurements data showing the abdomen with tiny details in such a high resolution that it is too beneficial in clinical settings and research for discovering abdominal features [6].

Web/mobile applications in health care have been extended that enable patients to send their signs and symptoms to the provider; those applications contain essential diseases, first aid, types of drugs, and also direct the patient to the specialist [7]. Health care system collects real-time bio-medical signals (e.g., ECG, pulse oximetry, and blood pressure) in different places on mobiles, a healthcare application is installed, and health data are coordinated for analysis and storage by a cloud computing system [8] in health care; big data can be characterized with the assistance of progressed information technology which observes information to make policy-making better; and a life chart can be used to research medical expenses and population aging, which applies evidence of policy-making [9]. Health care costs will be raised with the aging population; KSA has begun using big data technologies for approaching and managing elderly persons, and big data analytics is used to attain information from complex and enormous datasets obtained from data removal [10].

The current review provides a brief analysis of some productive efforts. In addition to the disadvantages and advantages of these technologies, privacy and security have been discussed in phases of big data analytics in healthcare. Big data analytics has linked the difference between organized and unstructured data. The transition to an integrated data environment is a recognized obstacle to overcome. Big data's objective and guiding concept is to gather more information and more insights from this information and has the capacity to forecast future occurrences. Several reputable healthcare firms expect a robust growth rate in the healthcare data sector.

Literature Review of Healthcare Data

With the rapid digital transformation needs of the country, the interest and investment in health information technology (HIT) by hospitals have been expanded. Demand for Health Informatics (HI) graduates coming from different disciplines (e.g., healthcare professions, Information Technology, etc.) has also increased. Those graduates who want to practice healthcare in Saudi Arabia must obtain their professional classification and registration from the Saudi Commission for Health Specialties (SCFHS) in order to be eligible to work as health informatics in the Saudi healthcare sector [11].

The current SCFHS system for licensure and ranking of HI professionals was in need of an update due to the changing landscape of the profession and diversity of the applicants seeking HI licensure. The SCFHS ranking system is now based on the graduate's certification where applicants are placed in one of the four licensure tracks: 'technician' for healthcare practitioners who have a diploma; 'specialist' for healthcare practitioners with a bachelor degree; 'senior specialist' for healthcare practitioners with a master degree; and 'consultant' for healthcare practitioners with PhD or equivalent [11].

Multiple forms of healthcare data include biomedical signals, genomic data, sensing data, biomedical images, and social media [12]. Genomic data analysis lets someone realize more about genetic markers, disease condition, consanguinity, and mutations; clinical text mining converts data from practical medical notes from disorganized format to applicable information, extraction of information, and natural language processing which extract helpful information from the massive volume of practical text. Social network analytics such as Web logs, Twitter, Facebook, social networking sites, and search engines helps to discover new health methods and worldwide health issues and trends based on different social media sources before analyzing the severity of the disease; therefore, reasonable diagnostic patterns should be used [13].

Table (1) represents the diagnostic plan for a definite diagnosis of the disease; the diagnosis is based on three conditions (frequency, pattern, and scale) [14].

Table (1): The scheme for diagnosis of diseases in a system

Disease and illnesses	Diagnostic method bases	Health measurements
Diabetes mellitus	Scale-based, frequency-based	**Blood glucose**
Pulmonary disease	Scale-based, frequency-based	**Oxygen saturation**
Cardiac disease	Pattern matching, frequency	**ECG**
Infectious disease	Frequency-based, scale-based	**Temperature**
Hypotensive disease	Frequency-based, scale-based	**Arterial pressure**
Gastrointestinal tumors	**Frequency-based**	**Video capsule pill**

Figure (1) presents the conceptual framework and approach. VHC is designed to have services like Virtual consultations (VC), Tele-pharmacy (TP), Virtual storage space (VSS), and Virtual Community (VCom). Virtual consultations would have two phases: In the first phase appointments/consultations would be conducted via videoconferencing, and in the second phase, chat sessions or off-schedule consultations are considered. The EHR is made available to health professionals (HP) and patients during any of these sessions. In addition to the VC, patients may also have an in-person consultation with the HP. At the end of the appointment, the patient and the HP could set a date and time for the next session. The same approach is followed to support the Mobile assisted health care system using smartphones. The focus here is to aid the senior citizens in sticking to their medication regimes and sending reminders so that they do not skip the prescribed doses [15].

This can also be used to inform the connected pharmacy to supply refills when low stock levels are reached. Another service included in the VHC is tele-pharmacy (TP), which allows the pharmacist to accept electronic prescriptions and transmit drugs to the patient's house via courier. Patients can use TP to document their treatment history on charts and consult on available medications. The system includes a Virtual Storage Space (VSS) to hold verified information about numerous diseases, which links to other web pages for patients and HP. All links are classified according to their source and organized into categories. The virtual community (VCom) is where people may share illness information, raise awareness about various diseases, and express their thoughts or comments on articles and news items [15].

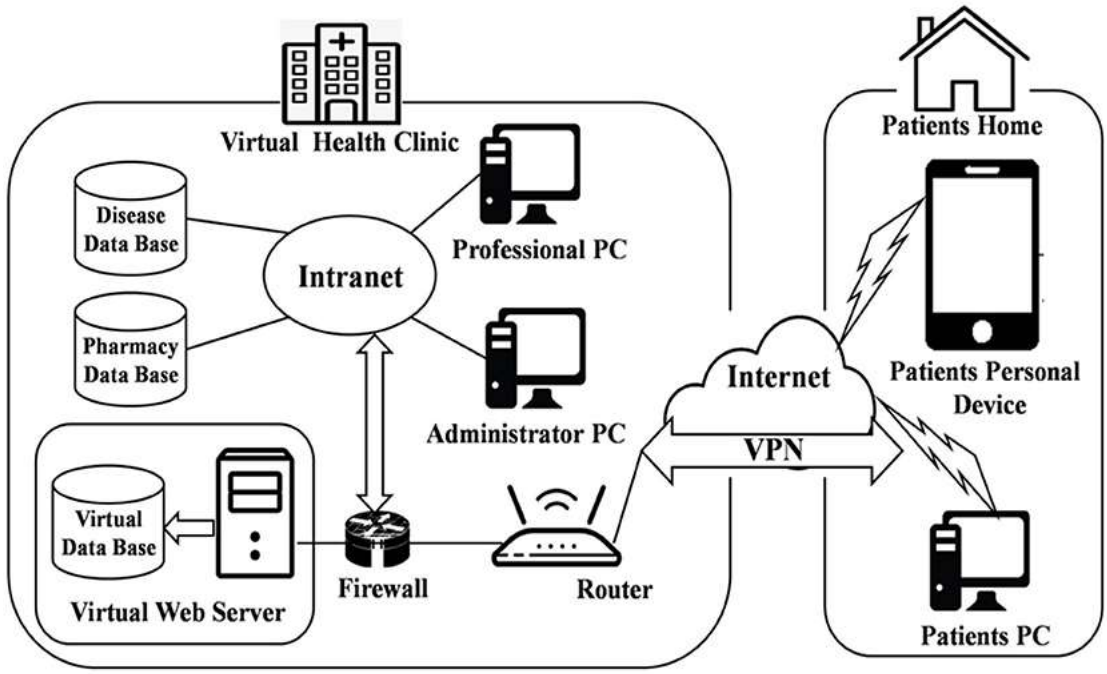

Figure (1): the conceptual framework and approach of digitized health- related information

This is about wearable functions of the physiological sensors, and then referred to as mobile physiological sensor systems, designed for gathering user information through different sensors. These sensors measure a patient's vital signs, including ECG, temperature, oxygen saturation, pulse rate, and blood pressure. After that, those real-time data will be shown on the user's smartphone and sent to the health care cloud. Cloud systems can analyze and make classification using machine learning methods and store information in a private and secured manner [16].

Healthcare data has been extended with a continuous stream of recent data elements and relationships, various data ranges from individual health information to epigenomes, and copious integration approaches are accepted, such as view integration (emerging and bringing together various databases), link integration (presentation in a web page), warehousing (setting data into a common database), and mash-ups (making a new web application from more than one web- based resource). All these methods make joining data flexible in many ways across sources but nevertheless consist of inadequate computable joined data or integration [17].

Healthcare Informatics and Analytics (HCI&A)

With the extensive adoption of database methods by different healthcare settings, HCI&A arose from the context of data management and analytics [17]. The goal during that time was to accomplish three goals: raise income, increase employee productivity, and save money [17]. These data management systems largely depend on technology for collecting, extracting, and analyzing health data. From a data-centric perspective, HCI&A may be compared to HCI&A, in which data is wholly organized, homogeneous, and stored in relational database management systems (RDBMS). In addition, three other significant factors contributed to the medical domain's artificial intelligence and data analytics: the medical domain, the web, and data. It consists primarily of Web technologies, Health apps, services, tools, and Medicine solutions [17].

A hospital or healthcare institution distributes content on Web without interacting with patients; it is primarily an online content repository. In the context of healthcare, Web aims to create an online presence for healthcare providers that make their information available at any time to all clients (primarily patients). In its cover of Web tools and methods, HCI&A encompasses the fundamentals of web technologies (HTML and HTTP), emerging web technologies (XML), and hypertext. Consumers and service providers cannot be involved in HCI&A technologies. Provider-centric approaches are at the center of Medicine and Health. Database technologies such as ware-houses are used to integrate healthcare data management systems [17].

Various statistical tools and data mining tools are also available in HCI&A to classify, segment, cluster, and analyze health data. The leading commercial healthcare informatics systems from IBM, Oracle, and Microsoft already incorporate some HCI&A features. In addition to extracting, transforming, and loading data, we also have database querying, data mining, and visualization packages within HCI&A. However, the software must also be able to perform some intelligence and analytical tasks [17].

Big Data Analytics in Healthcare System

Healthcare data digitization is the result of big data development and revolution. The rapid growth in data over the past few years led to the announcement of a new domain called big data. In information technology, the term “big data” is usually used to express enormous data that are too big and hard to deal with by the traditional database [18]. Intelligent healthcare systems, including big data analytics, make new and mobile health, saving medical costs and expanding efficiency [19]. Predicting pharmaceutical outcomes by predictive analytics, people who get the most benefit from pharmaceutical interventions are recognized, making pharmacists understand more about the side effects and risks of the medications [20].

Handling precision medicine is done by data collecting and management (sharing data, storing data, and privacy) to analytics (data merging, data processing, and visualization); compound and complex biomedical data which are enormous are becoming accessible due to biotechnologies progression, and analytics of big data is acquired to use these different data. It covers application sectors such as imaging, health, sensor, and bioinformatics [21]. For big data analytics, accuracy is essential; personal health records (PHRs) may contain typing errors, abbreviations, and mysterious notes; medical personal data input may contain errors, or it may be put in the wrong environment, which affects the efficacy of the collected data instead of getting uploaded by the professional trainee and medical practitioner in a clinical environment, and gathering data from social media may result in inaccurate prediction [22].

Fast-growing noise data is a significant problem; heterogeneous results are caused by various degrees of quality and completeness, which leads to false discoveries; there are two main problems, which are the inappropriate quality of data and biases because of absent randomization; big data value elevation are made by connecting various and analyzing all existing data [23]. Big data depending systems have progressed, including patient discharging records, electronic certificates of death, and medical claim data, which use the coding of International Classification of Diseases (ICD), and using big data courses in strategies of surveillance from the internet and social media has been preferred [24]. Data technologies like SQL databases have established healthcare processing. Some features like rational relationships and local access between logical and physical data spreading are significant to upgrading and performing parallel processing in database distribution [25].

Clinical and molecular information has been proposed in a big data-driven approach. Therapeutic medications and biomarkers are spotted in the approach. Following preclinical or clinical accuracy is accomplished by cross-species analysis; hence, the cost and time of biomarkers and therapeutics are decreased [26]. The primary function of the warehouse is for structured data and has a set of modules for unstructured data analysis. Initial accomplishments of substructures or frameworks were built for a big data paradigm. The framework used a Hadoop cluster for running modules, and distributed counting ability is used in big data according to research [27].

Utilization of the enormous data storage and reliability of the Hadoop big data by the system makes a considerable reduction in storage and upgrade costs. Mobile applications are widespread, keeping doctors and patient users in touch, decreasing complex medical communications and increasing digitalization. Hadoop is a software framework that uses a master-slave. A group of essential background programs is mandatory to get Hadoop running in a completed cluster softly; it is also spread by a large amount of data and progress by Apache Foundation; it is likely to develop a distributed program is capable of dividing a large amount of program into small working units, making the cluster's ability to make high-speed storage [28].

Map Reduce is based on rough set theory RST which is used for reducing attributes and includes these procedures for characteristic acquisition and accomplishes them on the Map Reduce parallel large-scale rough set method which is used in runtime systems like the Phoenix, Twister, and Hadoop to get features from the big database by data mining two acceleration of computation of equivalence classes done by using the framework structure of the (key, value) pair; Map Reduce parallelizes traditional attribute reduction [29].

Industry precision medicine is a sort of big data application in health issues, including the manufacturing of medical drugs and devices; it is considered as a strategic plan; this application benefits from IoT, industry, and multi topic. It has been suggested that it makes sense of big data with artificial intelligence, next-generation technology, and IoT [30]. Based on IoT technology, an intelligent healthcare framework has progressed for anyone during workouts; the Bayesian belief network uses an artificial neural network model to predict a patient's health-related susceptibility. There are four critical areas of big data analytics: model development, business models, data management, and visualization [14].

Challenges of Big Data in Healthcare Systems

Big data has been evolving, introducing challenges and problems caused by the exponential growth of healthcare data. The constant changes of big data present many challenges in analyzing, storing, and recovering huge amounts of data. Conventional or standard database systems cannot be used to process, store, and take information due to their massive and enormous volume [31]. Big data issues that generally happen in healthcare organizations are covered by **four main categories** [32, 33]: **a huge amount of unstructured data** are included in big clinical data like handwritten data and natural language, **a reasonable degree of difficulty** is brought by clinical big data's analysis, **integration, and storage**.

It is insufficient for agencies to share structured data, and unstructured data sharing among organizations is more complicated. It is a great challenge how to effectively mine an enormous amount of unstructured data. Big data has some characteristics. One of them is variability in data sources; medical data has potent timeliness; having appropriate moments of medical care is an example. In the medical industry, data processing speed is in great demand particularly while patients' situation deteriorates quickly. The data privacy and security of the patients and ill persons are influenced by challenges and disputes with these real-time applications like cloud computing used to analyze data [33].

Recently cloud computing has offered new possibilities for medical big data mining and sharing. Before cloud computing can become even more practical, several challenges must be overcome [34]. First, cloud computing offers a simple and flexible way to mine resources. However, it elevates the risk of privacy disclosure. It is a fact that is clinically evident in clinical informatics. Second, importing or exporting an enormous amount of data in medicine to the cloud (petabyte). Network bandwidth increases the cost of data and restricts the speed [35].

Economic Challenges: The medical field facilities of patients and health care providers such as doctors are dependent on paid services. It disproportionately negatively impacts technology advancements in connection with this process [36]. Big data technology challenges being highly fragmented and enormous, big data in healthcare leads to information quality and technology problems, making it a barrier to accomplishing healthcare vision [37]. Security and privacy issues along with the history of big data include the privacy of healthcare data which is serious because of potentially essential and sensitive information about individual healthcare providers. In order to make healthcare data unavailable in public, it must be secured from unauthorized access, preventing the data from attackers. This means security is the most important task, which is also a challenge [38].

Privacy: The most vital fault is the lack of intimacy and privacy. Big data must have access to almost everything, even social media life and private recordings, to have enough effects. However, because of revealing private information, the price is paid. Moreover, there is no patient freedom. However, there are regulations for stating medical recordings' privacy. However, they are not considered since it is believed that the information of someone should not be forbidden. At the same time, it is related to their health. The privacy risks associated with big data in health care have been stopped in articles such as big data privacy and security in health care [39].

Health Information Systems on the Cloud: The adoption of cloud-based platforms has improved and streamlined the design, development, and deployment of clinical information systems, hospital administrative information, and medical images [19]. Several such structures are in place to facilitate data collection (for example, the entities are often provided with mobile user interfaces to cloud services to gather and manage healthcare information). In addition, these systems facilitate information exchange between various medical structures, hospitals, and patients since they integrate data in several ways. The performance of the system is rarely considered. Security and privacy, considered essential, are often at the center of their design [19].

Tele-pathology, Tele-health, and Disease Surveillance: Tele-pathology services were envisioned as a possible outcome of combining robotic microscopy, video imaging, databases, and the then new availability of broadband telecommunications. Many contributions demonstrating ICT applications have been presented, illustrating how ICTs can assist with telemedicine, tele-pathology, and disease monitoring. Research has been conducted on two problems: (1) general frameworks for most cases and (2) studies that focus on particular diseases, such as cancer detection, cardiovascular disorders, diabetes, Parkinson's disease, and Alzheimer's disease. These monitoring systems may then be utilized as a tool for large-scale research and as a means for customizing therapies [31].

Likewise, surgery is expected to become more transparent in the future. Open surgery operating rooms often use video cameras for lighting. It allows an infinite number of viewers to view the surgical operation. Tele-consultation is possible with these instruments, eliminating the need for the consultant to be physically present. A remote consultant may use tele-presence during surgery if an active camera holder is used and the remote consultant can move the camera. It is physically impossible for the surgeon to see the operating room when tele-surgery is used. The availability of limited virtual pathways to fog services at the edge could assist in closing the gap when best effort internet connections are insufficient for some types of applications (e.g., to recreate the effect of a microscope locally). Providing remote federated sites with tools for offloading sophisticated image processing and data mining operations, it may, for instance, allow remote federated sites to cooperate on nontrivial diagnoses without experiencing increased cloud access latency [32, 33].

Big Data Management in the Healthcare System

Healthcare activities generate large amounts of data. Analytical procedures should be used to derive actionable judgments from data management technologies. This section is divided into five subsections: machine learning-based, agent-based, cloud-based, heuristic-based, and hybrid-based. Further, in this section, the chosen articles are described in their approach, differences, advantages, and drawbacks. This section examines the most common machine learning techniques for managing extensive healthcare data along with their fundamental characteristics [35]. In the last few years, machine learning methods have been used to process large amounts of data based on artificial intelligence methods and historical databases. Therefore, machine learning techniques can be compellingly applied to this problem [35, 36]. Machine learning algorithms play a significant role in managing massive biomedical data based on current issues in biomedical data [36].

Discussion on Intelligent Health Care

Sensor data are primarily unstructured in intelligent health care. Sensor-based health and wellness monitoring generate unstructured data beyond the human ability to process and analyze manually. There is a huge gap between the potential and the utility of such an enormous amount of unstructured data. The vast amount of unstructured data from streaming sensors is useless due to its variability and complexity. Data analytics pipelines for intelligent healthcare applications follow a similar process to the standard analytical method. Data management, processing, and finding are critically important in health care [(37, 38)]. The correct data must be collected at the right time and context for an effective data discovery process. There needs to be an end to the division between numerous fields, such as medical science and computer science, for context-awareness in healthcare applications [(39)]. Data duration is, therefore, more useful when addressing effective data discovery when it comes to improving knowledge of patient physiological and psychological care.

Interpretation of Data: Predictive analytics may be more effective when combined with structured and unstructured EHR data. Clinical events can be extracted from EHR data, and comparable phrases can be categorized in semantic space. By concatenating their representation using semantic space, structured and unstructured data are integrated more efficiently than if the occurrences are represented separately. Using semantic spaces to extract clinical language from EHR, diverse and distributed representations can predict clinical outcomes effectively. The lack of an agreed-upon standard for terms, acronyms, and abbreviations further complicates the semantic categorization of datasets [(40)].

This factor may impair the effectiveness of semantic classification based on similar terms. Various types of information can be collected from health records for purposes such as pharm-acovigilance, phenol-typing, and illness detection. Data from EHRs, EMRs, PHRs, and omits provide a wealth of information for many different medical fields. However, they should also be used to enhance healthcare. The model has been evaluated through interviews with domain experts, following the combination of clinical and genomic data for deep cancer phenol-typing. In this study, real-time datasets could neither be used to assess the representation standard nor assess the suggested model. A robust knowledge base and accurate data modeling may facilitate using unstructured clinical notes from multiple institutions. The interpretation of data is equally important as obtaining usable information from various forms of health records, and this is called enhanced unstructured data analytics [(41)].

Quality of Data: The literature has identified that several quality parameters can be used to enhance and assess significant data quality, such as correctness, completeness, consistency, timeliness, objectivity, interpretability, and accessibility. Unstructured, heterogeneous, and noisy data add to the difficulty of this task because of their heterogeneity, lack of structure, noise, and the lack of a preset model [(42)]. In addition to understanding psychological disorders, social media analytics helps to understand society's most prevalent illnesses. Social media analytics has most of the quality issues compared to other fields because postings, reviews, and comments cannot be standardized. Several linguistic issues impede clean analytics. It may be possible to increase analytical efficiency by using hash tags. However, computer, media, and healthcare knowledge are necessary to understand healthcare social media better. As part of effective healthcare analytics, database aggregation and data cleansing may reduce data heterogeneity, lack of structure, and other quality challenges [(40, 41)].

Conclusion

This paper is a brief discussion of some successful work. Privacy and security have also been presented in phases of big data analytics along with the faults and benefits of these technologies in big healthcare data. Big data analytics has held the gap between structured and unstructured data. A well-known obstacle to overcome is the shift to an integrated data environment. The aim and principle of big data are gaining more information; more insights from this information and the ability to predict future data healthcare market show a rapid growth rate which several reliable healthcare companies project. However, in a short time, we have seen a range of analytics in use which has shown improvement effects on health care industry decisions. Computational experts have been forced by the exponential growth of medical data from different domains to design strategies to interpret and analyze various amounts of data. In every area, big data challenges are as follows: storing, searching, capturing, sharing, and analyzing data. Some extra challenges include real-time processing, data quality, privacy and security, and heterogeneous data. Also, healthcare data standards are among the challenges of big data analytics in healthcare systems [(43)].

References

1. Alotaibi, A., Saleh, W., Abdulbaqi, A., & Alosaimi, M. (2022). Health Research Priority Agenda for Ministry of Health, Kingdom of Saudi Arabia from 2020 to 2025. Journal of Epidemiology and Global Health, 12(4), 413-429.
2. Bhavnani, S. P., Narula, J., & Sengupta, P. P. (2016). Mobile technology and the digitization of healthcare. European heart journal, 37(18), 1428-38.
3. Hemingway H, Asselbergs FW, Danesh J, Dobson R, Maniadakis N, Maggioni A, et al. Big data from electronic health records for early and late translational cardiovascular research: challenges and potential. Eur Heart J. 2018; 39(16): 1481–95. https://doi.org/10.1093/eurheartj/ehx487
4. Price, W. N. BIG DATA, PATENTS, AND THE FUTURE OF MEDICINE, SSRN, New York, NY, USA, 2016.
5. Zhang, Y., Qiu, M., Tsai, C. W., Hassan, M. M. and Alamri, A. Health-CPS: healthcare cyber-physical system Assisted by cloud and big data. IEEE Systems Journal, vol. 11, no. 1, pp. 88–95, 2015.
6. Ulfarsson, M. O., Palsson, F. Sigurdsson, J. and Sveinsson, J. R. Classification of big data with application to imaging genetics. Proceedings of the IEEE, vol. 104, no. 11, pp. 2137–2154, 2016.
7. Panda, M., Ali, S. M., and Panda, S. K. Big Data in Health Care: A Mobile Based Solution, in Proceedings of the 2017 International Conference on Big Data Analytics and Computational Intelligence (ICBDAC), Chirala, Andhra Pradesh, India, March 2017.
8. Tawalbeh, L. A., Mehmood, R., Benkhlifa, E., and Song, H. Mobile cloud computing model and big data analysis for healthcare applications, IEEE Access, vol. 4, pp. 6171–6180, 2016.
9. Wang, C., Li, F. Wang L. et al., The impact of population aging on medical expenses: a big data study based on the life table, BioScience Trends, vol. 11, no. 6, pp. 619–631, 2017.
10. Tsuji, Y. Medical privacy issues in ageing Japan, Australian Journal of Adult Learning, vol. 18, no. 1, 2017.
11. Almalki, M., Jamal, A., Househ, M., Alhefzi, M. A. (2021). multi-perspective approach to developing the Saudi Health Informatics Competency Framework. International Journal of Medical InformaticsVolume 146, February 2021, 104362
12. Ta, V.D. Liu, C.M., and Nkabinde, G. W. Big data stream computing in healthcare real-time analytics," in Proceedings of the 2016 IEEE International Conference on Cloud Computing and Big Data Analysis (ICCCBDA), Chengdu, China, July 2016.
13. Al, T. E. Big data stream computing in healthcare real-time analytics, IEEE International Conference, 2016.
14. Verma P., and Sood, S. K. Cloud-centric IoT based disease diagnosis healthcare framework, Journal of Parallel and Distributed Computing, vol. 116, pp. 27–38, 2018.
15. Mendelson, D. Legal protections for personal health information in the age of big data: a proposal for regulatory framework, ETHICS, MEDICINE AND PUBLIC HEALTH, vol. 3, no. 1, pp. 37–55, 2017.
16. Leu, F.Y., Ko, C.Y. You, I., Choo, K. K. R. and Ho, C.L. A smartphone-based wearable sensors for monitoring real-time physiological data," Computers & Electrical Engineering, vol. 65, pp. 376–392, 2018.

17. Murphy, S. N., Avillach, P., Bellazzi, R. et al., “Combining clinical and genomics queries using i2b2 – three methods,” PLoS One, vol. 12, no. 4, Article ID e0172187, 2017.

18. Raja, R., Mukherjee, I. and Sarkar, B. K. A Systematic Review of Healthcare Big Data, Scientific Programming, vol. 2020, pp. 1–15, Article ID 5471849, 2020.

19. Pramanik, M. I., Lau, R. Y., Demirkan, H., and Azad, M. A. K. Smart health: big data enabled health paradigm within smart cities, Expert Systems with Applications, vol. 87, pp. 370–383, 2017.

20. Hernandez, Y. Zhang, and September, “Using predictive analytics and big data to optimize pharmaceutical outcomes,” American Journal of Health-System Pharmacy, vol. 74, no. 18, pp. 1494–1500, 2017.

21. Wu, P. Y., Cheng, C. W., Kaddi, C. D., Venugopalan, J. Hoffman, R., and Wang, M. D. Omic and electronic health record big data analytics for precision medicine,” IEEE Transactions on Biomedical Engineering, vol. 64, no. 2, pp. 263–273, 2017.

22. Andreu-Perez, J., Poon, C. C. Y., Merrifield, R. D., Wong, S. T. C., and Yang, G. Z. Big data for health, IEEE J Biomed Health Inform, vol. 19, no. 4, pp. 1193–1208, 2015.

23. Sacrista´n and T. Dilla, J. A. No big data without small data: learning health care systems begin and end with the individual patient, Journal of Evaluation in Clinical Practice, vol. 21, no. 6, pp. 1014–1017, 2015.

24. Simonsen, L., Gog, J. R., Olson, D., and Viboud, C. Infectious disease surveillance in the big data era: towards faster and locally relevant systems, Journal of Infectious Diseases, vol. 214, no. 4, pp. S380–S385, 2016.

25. Salavati, H., Gandomani, T. J., and Sadeghi, R. A robust software architecture based on distributed systems in big data healthcare, in Proceedings of the International Conference on Advances in Computing, Communications and Informatics, pp. 1701–1705, Udupi, India, September 2017.

26. Wooden, B., Goossens, N., Hoshida, Y., and Friedman, S. L. Using big data to discover diagnostics and therapeutics for gastrointestinal and liver diseases, Gastroenterology, vol. 152, no. 1, pp. 53–67.e3, 2017.

27. Istephan S. and Siadat, M. R. Extensible query framework for unstructured medical data -- A big data approach, in Proceedings of the 2015 IEEE International Conference on Data Mining Workshop (ICDMW), pp. 455–462, Atlantic City, NJ, USA, November 2015.

28. Zhang X., and Wang, Y. Research on intelligent medical big data system based on Hadoop and blockchain, EURASIP Journal on Wireless Communications and Networking, vol. 2021, no. 1, p. 7, 2021.

29. Ni, J., Chen, Y., Sha, J., and Zhang, M. Hadoop-Based Dis- tributed Computing Algorithms for Healthcare and Clinic Data Processing, in Proceedings of the 2015 Eighth Interna- tional Conference on Internet Computing for Science and Engineering (ICICSE), pp. 188–193, IEEE, Harbin, China, November 2015.

30. zdemir V. Ö and Hekim, N. Birth of industry 5.0: making sense of big data with artificial intelligence, “the internet of things and next generation technology policy,” OMICS: A Journal of Integrative Biology, vol. 22, no. 1, pp. 65–76, 2018.

31. Reddy R., and Suresh-Kumar, P. Predictive big data analytics in healthcare, in Proceedings of the Conference: 2016 Second International Conference on Computational Intelligence & Communication Technology (CICT), Ghaziabad, In- dia, February 2016.

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

33. Sarkar, B. K. Big Data for Secure Healthcare System: A Conceptual Design, Complex Intell. Syst, vol. 3, no. 2, pp. 133–151, 2017.

34. Chen, J., Qian, F. Yan, W., and Shen, B. Translational bio- medical informatics in the cloud: present and future, BioMed Research International, vol. 2013, p. 1, 2013.

35. Hong, L., Luo, M., Wang, R., Lu, P., Lu, W., and Lu, L. Big Data in Health Care: Applications and Challenges, Data and In- formation Management, vol. 0, no. 0, 2019.

36. Wang L., and Alexander, C. A. Big data analytics in healthcare systems, International Journal of Mathematical, Engineering and Management Sciences, vol. 4, no. 1, pp. 17–26, 2019.

37. Firouzi, F., Farahani, B., Ibrahim, M., and Chakrabarty, K. Keynote Paper: From EDA to IoT eHealth: Promises, Challenges, and Solutions, IEEE Transactions on Computer- Aided Design of Integrated Circuits and Systems, vol. 37, no. 12, pp. 2965–2978, 2018.

38. Alexandru, G., Radu, I. M., Bizon, M. L., and June, A. Big data in healthcare - opportunities and challenges, Informatica Economica, vol. 22, no. 2/2018, pp. 43–54, 2018.

39. Abouelmehdi, K., Beni-Hessane, A., and Khaloufi, H. Big healthcare data: preserving security and privacy, J Big Data, vol. 5, no. 1, 2018.

40. Chiu, C. C., Wu, C. M., Chien, T. N., Kao, L. J., Li, C., & Chu, C. M. (2023). Integrating Structured and Unstructured EHR Data for Predicting Mortality by Machine Learning and Latent Dirichlet Allocation Method. International journal of environmental research and public health, 20(5), 4340. https://doi.org/10.3390/ijerph20054340

41. Zhang, D., Yin, C., Zeng, J. et al. Combining structured and unstructured data for predictive models: a deep learning approach. BMC Med Inform Decis Mak 20, 280 (2020). https://doi.org/10.1186/s12911-020-01297-6

42. Taleb, I., Serhani, M.A., Bouhaddioui, C. et al. Big data quality framework: a holistic approach to continuous quality management. J Big Data 8, 76 (2021). https://doi.org/10.1186/s40537-021-00468-0

43. Dash, S., Shakyawar, S. K., Sharma, M., and Kaushik, S. Big data in healthcare: management, analysis and future prospects, Journal of Big Data, vol. 6, no. 1, 2019.