ISSN 2278-3091
Volume 8, No.6, November – December 2019
International Journal of Advanced Trends in Computer Science and Engineering
Available Online at http://www.warse.org/IJATCSE/static/pdf/file/ijatcse55862019.pdf
https://doi.org/10.30534/ijatcse/2019/55862019

Big Data Analytics for Value-Based Care: Challenges and Opportunities

Ibrahim A. Atoum[1], Nasser A. Al-Jarallah[2]
[1]Department of Computer Science and Information Systems AlMaarefa University, PO Box 71666, Riyadh 11597, Kingdom of Saudi Arabia, iotoum@mcst.edu.sa.
[2]Dean of Applied Science College at AlMaarefa University, PO Box 71666, Riyadh 11597 Kingdom of Saudi Arabia, njarallah@mcst.edu.sa.

ABSTRACT

Big data denotes large volume datasets that grows exponentially over time and cannot be processed or stored using traditional data management tools. There is a huge amount of new data that is stored daily in all areas of our life. This data needs to be collected and analyzed effectively to streamline the decision-making process. Big data analytics (BDA) are techniques that provide a way to analyze this huge quantity of data with the purpose of drawing conclusions about them by driving the business outcomes and real time decisions. With a wide range of data available in the healthcare sector including financial and clinical data, research and development, management and operational data and critical care units, big data in healthcare can generate meaningful insights to improve overall operational efficiency in the industry. There are various benefits for big data analytics in the health-care industry which encompasses advanced patient care; improved operational efficiency; evaluation of practitioner performance; reduced patient cost and researching cure for diseases. The main objective of this study is to outline the main challenges and opportunities of applying big data in health care and how it can help in transforming the healthcare industry towards the value-based care.

Key words: Big data, Analytic, Value-Based Care, Data Warehouse, HealthCare

1. INTRODUCTION

There is a large quantity of heterogeneous data collected from different sources in health care organizations like claims and cost data [1][2]. This huge amount of data needs suitable artificial intelligence (AI) tools to become more beneficial in real time. Big data analytics can help AI to discover valuable hidden patterns and additional information such as patient needs and desires and market trends that can help in developing patient-oriented decisions. In spite of big data has been used rapidly and effectively in different industries, its use in the health care industry has been lagging behind due to poor data quality and small and inconsistent dataset [3] [4]. Although the availability of Electronic Health Records (EHR) has led to an evolution in health information systems, particularly in improving the quality of patient's health care, it faces two main challenges: the increasing volume of electronic data and the urgent need to use data analytics to interpret the meaning of this vast amount of data and provides predictability [5]. EHR have played an important role; especially after they have been integrated with imaging informatics [6].

Big data like all forms of technology offers many benefits but its application faces many challenges and problems as it is still in its emerging stages for some areas and in addition to being a complex technique. Most of the existing challenges are identified after organizations begin moving towards the usage of its technologies. The most important challenge is the dramatic and significant increase in data, especially unstructured data [7][8] [9]. Furthermore; companies urgently need to process updated data along with stagnant data so that they can build a better vision and enhance their decision-making capabilities [10] [11]. Moreover; the diversity of data sources and their regulatory standards makes them suffer from security problems, all security practices must be provided for collecting, storing and retrieving data **[12]**. Also, there is a shortage of experienced people in the field of big data, which slows the transition to the application of its technologies [13]. It is necessary to overcome all of these challenges to get all the positive advantages of this technique in order to better understand customers and make better decisions.

Big data analytics can be used to develop and support value-based care. Value-based care or value – oriented health care is one of the healthcare delivery models [14] whereby payment to providers is paid according to the level of health outcomes of the patient instead of the delivered amount of health care services**.** This is due to the fact that **o**ver the recent period, consumer demand for premium health care has increased. Subsequently, the healthcare industry has moved from being a volume-based industry to value-based industry [15] [16], which require smarter decisions.

An analysis of 26 big data implementation cases shows how big data analytics are used as means of transforming information technology into value-based business [17]. The

healthcare industry is undergoing a major shift in the use of big data analytics which will be reflected on five different directions [18][19]: Value-based patient-centric care , the healthcare internet of things , reducing fraud, waste, and abuse , predictive analytics to improve outcomes , real-time monitoring of patients.

This study will introduce how big data analytics supports value-based care and it's organized as follows: big data architecture in the healthcare industry is discussed in Section Two. Section Three illustrates BDAs and Value-based Care, and conclusions are presented in Section Four.

2. BIG DATA ARCHITECTURE

Health data are huge with structured, unstructured and semi-structured data. It is practically impossible to process the data which has fixed schema (unstructured and semi-structured data) using traditional databases. Instead, to facilitate the analysis and categorization of this huge amount of data, the data is placed in a central repository called data warehouse or centralized hub [20] as shown in Figure 1.

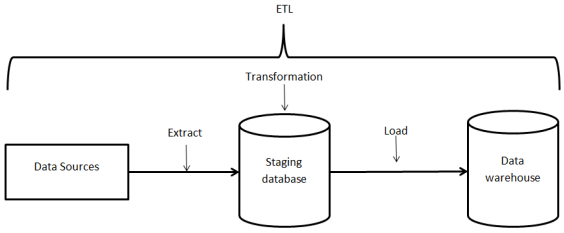

Figure 1: Data Warehouse

The data is loaded from different data sources into the staging area which is a temporary storage area responsible for extract, transform and data cleansing. Then it's moved to the data warehouse to be organized for query and analysis. A data warehouse is where an OnLine Analytical Processing (OLAP) database exists on top of another OnLine Transaction Processing (OLTP) databases as shown in Figure 2. The important fact is that the databases are not structured to do analytics while data warehouse are [21].

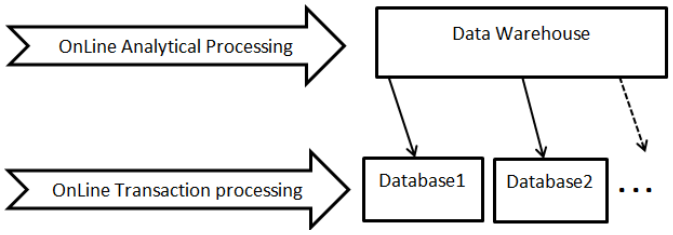

Figure 2: Data Warehouse vs. Databases

Data Warehouse is architecture of data repository that handles structured data, whereas Big Data is a technology to handle massive structured; unstructured and semi-structured data which constitutes the variety vector of big data. Big data is a modern technology used to launch a business analytics solution for growing data volumes and types stored in different applications with different formats. It is the overall system used by the organization for the purpose of big data analytics [22] and it is characterized by three key concepts: volume, variety, and velocity [10]. These three concepts have been introduced in a 2001 Meta Group research publication [23] [24]. Other concepts later attributed with big data are veracity and value. The size and diversity of data is the reason behind the organization needs of big data analytics.

Big data management architecture in healthcare industry has four main logical layers [25] as shown in Figure. 3. The layers are logical; so it doesn't mean that their functions are run on separate machines.

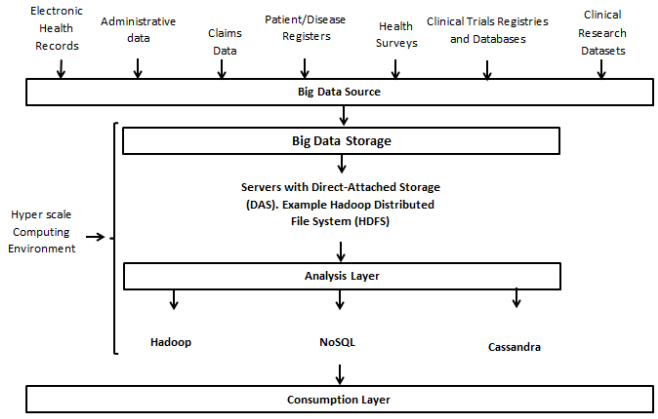

Figure 3: Big Data Architecture for healthcare Industry

The first layer represents the big data sources which may include EHR, administrative data; etc. The heart of this architecture is referred to as the hyper scale computing environment and it represents two logical layers. The first one is the big data storage layer which located between the big data sources and the analysis layer. In this layer, data is extracted from different data sources with different formats; stored in servers with direct-attached storage like HDFS and then it is converted to a format that is compatible for the analysis of the data [26]. The second layer is the analysis layer which receives data from the storage layer. This layer uses analytical tools like Hadoop, NoSQL, Cassandra, etc., to determine trends, as well as drawing their conclusions. The last layer is the consumption layer which consists of presenting the analysis results to human applications or business processes.

3. BDAS AND VALUE-BASED CARE

It has become apparent that health care cost has become very expensive, especially for poor and middle-income families, this prompting the motivation to think about changing the pay-for-care approaches from fee-for-service (FFS) to value-based care (VBC). In the FFS payment model, the payments for health services depend on increased patient visits and procedures while on VBC based on payment for the actual health outcomes. There are many benefits to the VBC model that covers all aspects of the

health care system which contains patients, providers, suppliers, payers and society as shown in Figure 4.

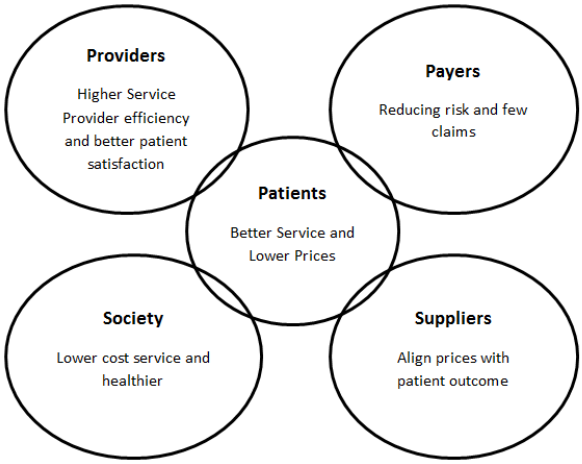

Figure 4: Components of Healthcare Systems.

BDA tools can examine a large number of variables in search for combinations that accurately predict results [27]. In the case of health care, these variables may be derived from millions of health care claims or EHRs and can be used by BDAs to correct erroneous diagnosis and control the progression of the disease. The most wasted money in health care is spent on misdiagnosis, so careful diagnosis is one of the first steps to getting more value-based care.

BDAs power data-driven transformations that seemingly combine financial and administrative data to aid patient care efforts, provide better services, and improve existing procedures [28]. The adoption of these tools over these data can reduce cost and develop treatment; it may generate a growing concern among service providers about security. This concern or challenge is acute when solutions are cloud-based, which means the ability to access confidential patient information. In the case of value-based care, instead of paying for each service or visit; payments are collected, frequent procedures and visits will be eliminated, patients will move more quickly and the number of visits to the hospital and to the emergency room will be lower thus the most important value-based consequence can be considered as reducing cost.

Treatment can be improved by adopting these tools through eliminating redundant tests, improving survival rates in the Intensive Care Unit (ICU) and promoting preventative care using wearables [21] [29]. Wearables used to detect some diseases which help users seek timely medical attention will improve the prevention services. In Reality**,** the most data-dependent clinical area is the critical care in all its forms which includes: Intensive Care Unit (ICU); Pediatric Intensive Care Unit (PICU), Neonatal Intensive Care Unit (NICU) or Surgical Intensive Care Unit (SICU). This high-variety, high volume and high-velocity data – centered areas require a medical decision support system under the umbrella of big data to drive the decision- making process as shown in Figure 5 which will inevitably increase the survival rates [30] [31]. Although there are many studies of applying big data analytics in this field, there are very few implemented examples which led to a lack of confidence by clinicians in these systems.

Some other benefits of big data which the industry has experienced are error minimization and precise treatments through the accurate and detailed knowledge of patients; for example the problem of dangerous high dose threshold levels can be solved by using big data. It can be done by reviewing medication ordering patterns to predict threshold and determine better parameters [32] [33]. This is one of the value-based health-care requirements of providing patients with better value rather than more visits, procedures and tests**.**

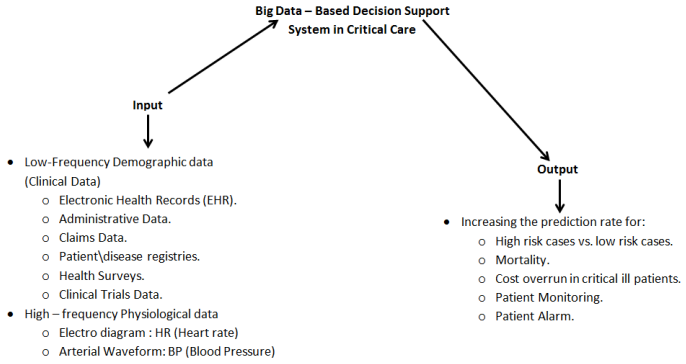

Figure 5: Critical Care BD Framework

Whenever health care is directed towards value-based care, reimbursement should be linked to the quality results due [11] [34]. First of all, data sets can be used for each patient through all available data sources. The patient's behavior, attitudes and habits can be understood and therefore result in sound decisions making in order to improve patient outcomes. In addition, they can be used to implement risk management programs by understanding the clinical and financial risks of the patient and therefore more targeted interventions can be conducted for risk sites. Moreover, they can focus in human and financial resources on areas with the greatest return on investment through predictability and identifying opportunities with the highest clinical or financial outcomes. Finally, they can benefit by improving cost and satisfaction and by avoiding expensive services.

Additionally BDAs can help the hospitals in tracking staffing metrics like hours; shifts; etc., that may translate into success levels which will provide high value care for their patients. Moreover enhance billing efforts by tracking the patient's financial payment record as it is possible to determine who has the ability to pay or not. This feature leads to communicate quickly with patients to ensure that payments are collected and to streamline the payment procedures. All the components of health systems shown in Figure 5 get benefit from these assistances which latterly will support value-based care**.** The transform from FFS to FBS may face many challenges, including reshaping the design and performance of the current systems particularly the financial and clinical quality measures [35]. However there is a need to rebuild the technology infrastructure which provides better means of communication

and information exchange [36]. Also there is a need to invest in the human capital in order to enhance human health and well-being [37].

4.CONCLUSION

Data is available more than ever before. Quantity and quality of healthcare data can be exploited to raise the quality of the healthcare industry. Big data is accelerating the movement toward value-based healthcare. An important aspect of big data analysis is to discover the true meaning of data by taking advantage of important parts and neglecting unnecessary parts. Data collection, regardless of its sources and format, can be stored, using servers with directly attached storage and analyzed, using different data analytics like Hadoop which can present the analysis results. Despite all the benefits of big data analytics in the healthcare industry, it still face various challenges like the fundamental need to develop the infrastructure to be able to collect data from its various sources, as well as access control and security. In this paper, the structure of the application of big data in the field of health care has been reviewed. Furthermore, the study shows how healthcare big data can help in enhancing the value-based care. In the future; the different formats of the big data sources in the healthcare organizations will be considered; their storage mechanisms and the application constraints. Furthermore, the future of big data in critical care will be studied.

REFERENCES

[1] Prabha Susy Mathew, Dr. Anitha S. Pillai (2015) 'Big Data Solutions in Healthcare: Problems and Perspectives', *IEEE Sponsored 2nd International Conference on Innovations in Information Embedded and Communication Systems ICIIECS'15*.

[2] NM Alotaibi, M Abdullah, H Mosli, (2019). Agent-based Big Data Mining. International Journal of Advanced Trends in Computer Science and Engineering, 8(1).pp. 245-252
https://doi.org/10.30534/ijatcse/2019/4481.12019

[3] McKinsey & Company. (2019). *the big-data revolution in US health care: Accelerating value and innovation*. [Online] Available at: https://www.mckinsey.com/industries/healthcare-systems-and-services/our-insights/the-big-data-revolution-in-us-health-care [Accessed 20 Apr. 2019].

[4] Dhindsa, K., Bhandari, M. and Sonnadara, R. (2018). What's holding up the big data revolution in healthcare?. *BMJ*, p.k5357.
https://doi.org/10.1136/bmj.k5357

[5] Feldman, K., Davis, D. and Chawla, N. (20 15). Scaling and contextualizing personalized healthcare: A case study of disease prediction algorithm integration. *Journal of Biomedical Informatics*, 57, pp.377-385.
https://doi.org/10.1016/j.jbi.2015.07.017

[6] W Cooper, James & Ebadollahi, Shahram & Eide, Ellen & Neti, C & Iyengar, Giridharan. (2019). A Medical Imaging Informatics Appliance.

[7] P., D. and Ahmed, K. (2016). A Survey on Big Data Analytics: Challenges, Open Research Issues and Tools. *International Journal of Advanced Computer Science and Applications*, 7(2).
https://doi.org/10.14569/IJACSA.2016.070267

[8] T. K. Das and P. M. Kumar (2013). Big data analytics: A framework for unstructured data analysis, International Journal of Engineering and Technology, 5(1) (2013), pp.153-156.

[9] T. K. Das, D. P. Acharjya and M. R. Patra, Opinion mining about a product by analyzing public tweets in twitter, *International Conference on Computer Communication and Informatics*, 2014.
https://doi.org/10.1109/ICCCI.2014.6921727

[10] S. Kaisler, F. Armour, J. A. Espinosa and W. Money, "Big Data: Issues and Challenges Moving Forward," *2013 46th Hawaii International Conference on System Sciences*, Wailea, Maui, HI, 2013, pp. 995-1004. doi: 10.1109/HICSS.2013.645

[11] Zheng, Z.,Wang, P., Liu, J., Sun, S. (2015). Real Time Big Data Processing Framework: Challenges and Solutions, *Applied Mathematics & Information Sciences*, 9(6):3169-3190

[12] Merelli, I., Pérez-Sánchez, H., Gesing, S. and D'Agostino, D. (2014). Managing, Analyzing, and Integrating Big Data in Medical Bioinformatics: Open Problems and Future Perspectives. *BioMed Research International*, 2014, pp.1-13.
https://doi.org/10.1155/2014/134023

[13] Yoo, C., Ramirez, L. and Liuzzi, J. (2014). Big Data Analysis Using Modern Statistical and Machine Learning Methods in Medicine. *International Neurourology Journal*, 18(2), p.50.
https://doi.org/10.5213/inj.2014.18.2.50

[14] Wilson, H., Gole, J., Mishra, B. and Mishra, J. (2016). Value based Healthcare. *Advances in Management*, 9(1).

[15] Cortada, J. W., Gordon, D., & Lenihan, B. (2012). The value of analytics in healthcare: from insights to outcomes. IBM Global Business Services, Life Sciences and Healthcare, Executive Report.

[16] H. Kupwade Patil and R. Seshadri, "Big Data Security and Privacy Issues in Healthcare," *2014 IEEE International Congress on Big Data*, Anchorage, AK, 2014, pp. 762-765. doi: 10.1109/BigData.Congress.2014.112.

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

[18] McDonald, C., 5. big data trends in healthcare for 2017. *MapR Guide to Big Data in Healthcare. Retrieved from https://mapr. com/blog/5-big-data-trends-healthcare-2017*.

[19] A. Maqsood, (2017). Study of Big Data: An Industrial Revolution Review of applications and challenges. International Journal of Advanced Trends in Computer Science and Engineering, 6(3), pp. 31-41.

[20] McAfee, Andrew & Brynjolfsson, Erik. (2012). Big Data: The Management Revolution. Harvard business review. 90. 60-6, 68, 128.

[21] S. Sagiroglu and D. Sinanc, "Big data: A review," *2013 International Conference on Collaboration Technologies and Systems (CTS)*, San Diego, CA, 2013, pp. 42-47. doi: 10.1109/CTS.2013.6567202.

[22] K. Bakshi, "Considerations for big data: Architecture and approach," *2012 IEEE Aerospace Conference*, Big Sky, MT, 2012, pp. 1-7.
doi: 10.1109/AERO.2012.6187357.
[23] Strategy, P.B., 2011. getting value from Big Data. *Gartner Group Press Release*, *7*.
[24] Y. Demchenko, C. de Laat and P. Membrey, "Defining architecture components of the Big Data Ecosystem," *2014 International Conference on Collaboration Technologies and Systems (CTS)*, Minneapolis, MN, 2014, pp. 104-112. doi: 10.1109/CTS.2014.6867550.
[25] Hu, H., Wen, Y., Gao, Y., Chua, T. and Li, X. (2015). Toward an SDN-enabled big data platform for social TV analytics. *IEEE Network*, 29(5), pp.43-49.
https://doi.org/10.1109/MNET.2015.7293304
[26] D Ahamad, M Akhtar, S Hameed, (2019). A Review and Analysis of Big Data and MapReduce. International Journal of Advanced Trends in Computer Science and Engineering, 8(1). pp. 1-3.
https://doi.org/10.30534/ijatcse/2019/01812019
[27] D. J. Power, C. Heavin, J. McDermott, M. Daly. (2018) Defining business analytics: an empirical approach. *Journal of Business Analytics* 1:1, pages 40-53.
https://doi.org/10.1080/2573234X.2018.1507605
[28] Lee, J., Kao, H. and Yang, S. (2014). Service Innovation and Smart Analytics for Industry 4.0 and Big Data Environment. *Procedia CIRP*, 16, pp.3-8.
https://doi.org/10.1016/j.procir.2014.02.001
[29] H. Kupwade Patil and R. Seshadri, "Big Data Security and Privacy Issues in Healthcare," *2014 IEEE International Congress on Big Data*, Anchorage, AK, 2014, pp. 762-765. doi: 10.1109/BigData.Congress.2014.112.
[30] Sanchez-Pinto, L. N., Luo, Y., & Churpek, M. M. (2018). *Big Data and Data Science in Critical Care. Chest.*doi:10.1016/j.chest.2018.04.037.
[31] Paula Subias and Vicent Ribas (2018). Big Data for Critical Care. *Big Data CoE, 2018.*
[32] Higginson D. Phoenix Children's CIO says big data tactics reduce medical errors. *Healthcare IT News*. Jan 17, 2017.
[33] Balasuriya L ; Vyles D ; Bakerman P; et al.(2014). Computerized dose range checking using hard and soft stop alerts reduces prescribing errors in a pediatric intensive care unit. J Patient Saf. ; 13: 144-148
[34] Kansagra, A., Yu, J., Chatterjee, A., Lenchik, L., Chow, D., Prater, A., Yeh, J., Doshi, A., Hawkins, C., Heilbrun, M., Smith, S., Oselkin, M., Gupta, P. and Ali, S. (2016). Big Data and the Future of Radiology Informatics. *Academic Radiology*, 23(1), pp.30-42.
[35] Porter, M.E. and Kaplan, R.S., 2016. How to pay for health care. *Harv Bus Rev*, *94*(7-8), pp.88-98.
[36] D. J. Power, C. Heavin, J. McDermott, M. Daly. (2018) Defining business analytics: an empirical approach. *Journal of Business Analytics* 1:1, pages 40-53.
https://doi.org/10.1080/2573234X.2018.1507605
[37] De Mauro, A., Greco, M., Grimaldi, M. and Ritala, P. (2018). Human resources for Big Data professions: A systematic classification of job roles and required skill sets. *Information Processing & Management*, 54(5), pp.807-817.
https://doi.org/10.1016/j.ipm.2017.05.004