International Journal of Engineering Applied Sciences and Technology, 2019
Vol. 4, Issue 5, ISSN No. 2455-2143, Pages 64-73
Published Online September 2019 in IJEAST (http://www.ijeast.com)

USABILITY OF BIG DATA ANALYTICS WITHIN CLINICAL DECISION SUPPORT SYSTEMS

Christopher Otieno Okal
Department of Information Technology
Mount Kenya University, Nairobi, Kenya

Harriet Tsinale Loice
Department of Information Technology
Mount Kenya University, Nairobi, Kenya

Abstract - The adoption of electronic health record systems and other digital technologies such as Magnetic Resonance Imaging (MRI) techniques, automated laboratory tests, and body sensors have brought the era of big data technology into the healthcare industry. The use of big data technologies has the potential to provide medical organizations with powerful tools to gather and analyze large data volumes and to use this information to their advantage. However, special skills, systems, and capabilities are required to be able to analyze and extract useful information from big data. The objective of this paper was to explore the literature regarding the usability of big data analytics in supporting medical decision making. This information will guide healthcare organizations in understanding how they can adopt the utilization of big data to enhance decision making. A systematic review of evidence-based research articles from within the past five years was used to gather information in regards to this topic. The articles were derived from scientific databases. Based on the literature review, big data and big data analytics has the capability to improve decision making in the healthcare sector, predict disease outbreaks as well as the trends and patterns of the spread of such diseases, predict occurrence of medical phenomenon's such as hospital readmission, reoccurrence of diseases, and risk of infection among others. Moreover, big data analytics has the capability to help healthcare organizations to streamline processes within the healthcare setting. However, the process of integrating big data analytics in the healthcare setting follows distinct phases. Healthcare organizations also have to consider the challenges associated with adopting big data analytics. Nevertheless, based on the literature, big data analytics has the capability to improve delivery systems and outcomes within the healthcare sector.

Keywords: ***Big Data, Big Data Analytics, Medical Decision Making, Healthcare Sector, Data Acquisition, Data Storage, Data Management, Data Analytics, Data Visualization***.

I. INTRODUCTION

The value of big data has been optimized and appreciated in various sectors, such as banking and retail. However, the idea of big data in the healthcare industry is relatively new compared to other industries. Nevertheless, this concept is rapidly gaining popularity as most healthcare organizations and providers are gradually learning and appreciating the capability of big data in improving care delivery and patient outcomes.[1] There are various forces that have, for a long time limited the utilization of big data in the healthcare industry compared. The first major hindrance was due to the fact that the heartcare industry has been characterized by an environment where care providers make independent decisions and judgments. The conservative nature of this profession makes it challenging to implement new concepts such application of big data due to the resistance to change. Additionally, insufficient resources and lack of investment in information technology systems by a significant number of healthcare providers has also led to a reduction in the rate of utilization of big data in the healthcare industry.[2] The lack of sufficient training of individuals in the healthcare industry on how to integrate big data and big data technologies in the healthcare setting has also contributed to the low adoption of this concept in the sector. The lack of clearly defined procedures on how to apply big data as well as the security and privacy challenges associated with big data has also hindered its application.[3]

Nevertheless, the size of medical data has been on the rise in recent years. The size of medical data in 2012 was approximately 500 petabytes. This size is expected to increase to 25, 000 before 2020.[3] The data is also expected to increase in variety and complexity. The increase in size, variety, and

complexity of medical data has led most hospitals to begin making decisions based on this data. This implies that despite the late adoption of big data in the healthcare sector this data is essential for healthcare providers to improve care and service delivery towards their patient hence leading to better patient outcomes

II. STATEMENT OF THE PROBLEM

As stated before, the rate of adoption of Big has been slower in the healthcare industry relative to other industries such as banking. Despite the fact that utilization of big data has a significant potential to streamline and improve healthcare delivery in various ways, most healthcare institutions have been reluctant to invest in systems that increase their ability to utilize big data as well as big data analytics.

III. PURPOSE

Based on the above-stated problem the purpose of this paper is to explore the literature surrounding big data and its applicability in the healthcare industry hence building a case for the usability of big data and big data analytics. This information will play an important role in informing healthcare providers of the capabilities of big data and big data analytics, thus enabling them to increase their utilization in improving healthcare delivery and systems.

IV. METHODS

A literature review of scholarly sources related to healthcare big data and the applicability of big data in the healthcare industry was analyzed. These sources were drawn from various databases. Moreover, only journals and research articles from the past five years were included in the review.

V. LITERATURE REVIEW ON BIG DATA AND BIG DATA ANALYTICS IN HEALTHCARE

As stated before, the amount of data and information generated by the healthcare sector has been rapidly increasing in the recent past, and this trend is expected to continue into the future.[4] Historically, most of the records in the healthcare industry were stored as hard copies. However, current trends in digitization, which have been fueled by the increased in demand for quality healthcare services and the need to improve the efficiency of the healthcare delivery process have led the availability of data in soft-copy hence giving rise to the era of big data in the healthcare industry.[5]

There are various definitions that can be used to define big data. However, in the context of the healthcare industry, big data refers to huge sets of data that are complex to the extent that they cannot be managed using traditional data management systems nor can they be analyzed using the common data analysis softwares.[6] Raw big data may not be useful; however, this data can provide important insights, improve decision making, and service delivery if processed and analyzed. The complexity of big data management is as a result of its diversity and the speed in which it needs to be processed for it to have a significant impact.[7] There are various sources of big data in the healthcare setting these include notes from physicians, medical prescription records, imaging technologies, laboratory data, data from the pharmacy, insurance data, and other administrative and medical process records.[8] Additionally, big data in the healthcare environment may be generated from Electronic Health Records, medical sensors data, customers feedbacks, and posts on social media sites such as Facebook, Twitter, Instagram, websites activities and news feeds among others.[8]

To makes sense of data obtained from these sources, it is crucial for stakeholders in the healthcare sectors to be able to visualize this information in a format whereby they can observe trends, links, and patterns. These needs necessitate the concept of big data analytics which refers to the process of analyzing raw big data from different sources with the aim of portraying the data in a manner where an individual can observe the trends or patterns or draw insights to help in improving decision making and service delivery.

a) Characteristics of Big Data in the Healthcare Industry

There are multiple characteristics associated with big data in the healthcare industry. They include velocity, volume, variety, variability, veracity, and value.[9] The term volume implies that big data exists in the form of large data sets. On the other hand, the term velocity implies that big data is processed and it accumulates fast while the term variety implies that big data exists in different forms and can be obtained from different sources.[9] Furthermore, the term veracity refers to the degree of correctness or the ease with which the accuracy of data can be ascertained. However, big data has low veracity, which implies that its accuracy cannot be easily ascertained or validated, given the fact that accumulates fast and from different sources.[10] Variability implies that the nature and value of big data easily vary from one stage of its processing to the other. While the term value implies that big data is an important source of valuable information.[10]

b) Application of Big Data Analytics in the Healthcare sector

There are promising prospects for big data and big data analytics in the health care sector. The continued digitization of records in a healthcare institution is one of the core factors that has promoted the continued adoption of big data and big data analytics in the healthcare sector. Big data analytics is expected to improve healthcare delivery and outcomes in various ways.

The first major aspect of healthcare delivery that can be improved through the use of big data and big data analytics is supporting clinical decision making.[11] As stated before, Big data analytics has the capability to organize and learn from huge data sets such that the observable trends and patterns can be used to predict correct future outcomes hence improving the process of decision making which in most case is riddled by uncertainties about the future outcomes. Improved decision-making leads to better patient relationships as well as outcomes.[12] Improved and informed decision making also helps to reduce the chances for error.[13] Moreover, big data analytics has the capability of improving the provision of personalized care to patients by providing healthcare providers with real-time information about patients hence allowing them to understand a patient's condition and to engage in best practices when dealing with the identified problems.[14]

Similarly, big data analytics can also be used to visualize data on the spread and outbreak of disease and to predict future patterns in regards to such occurrences.[15] The ability to predict public health risks and diseases outbreaks makes it possible for stakeholders in the healthcare industry to undertake preventive and mitigative measures to protect or cushion the public from such occurrences.

Big data analytics also has the capability to improve clinical operations and care delivery systems. Studying client walk in trends, duration of service delivery, customer feedback among other process characteristics can play a significant role in aiding healthcare managers to understand how they can streamline or improve service delivery.[16] Moreover, the analysis of financial information can play a significant role in aiding hospital managers to ensure optimal resource allocation and in developing prudent financial policies and procedures.

VI. CASES OF SUCCESSFUL APPLICATION OF BIG DATA ANALYTICS IN THE HEALTHCARE INDUSTRY

Furthermore, research indicates that there are various cases of successful application of big data analytics in the healthcare sector in different regions and platforms around the world. One such case is the application of big data, and big data analytics to monitor the outbreak and spread of diseases. Google has successfully been able to predict disease outbreaks by analyzing search engine queries in different locations and their frequency.[17] An increase in the search frequency of a certain disease or its symptoms may imply an increase in the likelihood that there is a possible outbreak of the disease in a certain region. Research indicates that there is a significant increase in the demand for medical information and that most individuals are taking advantage of the widespread internet connectivity to meet their demand; this implies that there is a huge potential for other stakeholders in the healthcare industry to take advantage of social media and internet data to predict the outbreak of diseases.

In regards to the application of big data and big data analytics to support decision making, research shows that approximately 79% of healthcare organizations and care providers have adopted the use of informatics as a method of reducing the occurrence of medical errors.[18] Additionally, 61% of healthcare organizations in the United States are confident that big data and big data analytics has the capability to promote public health while 51% of the stakeholders believe that the application of big data in the healthcare industry has the capability to deal with the rising healthcare costs.[18] Big data analytics has the capability to help healthcare providers to streamline processes such as scheduling of activities, e-discharge, and e-referrals by using data and information from these activities to identify trends and patterns and to make improvements based on predictions.

Furthermore, big data analytics can also be applied to the advancement of evidence-based practices. Healthcare organizations are applying this capability of big data analytics to analyze the contents of the huge database of medical research and literature hence allowing physicians to be able to make more informed decisions.[19] Big data analytics can also help individuals to understand about their own conditions by presenting information in a manner that is easy to understand hence allowing such individuals to participate in decision making in regards to their health conditions.[18]

Furthermore, institutions such as Heritage Health have initiated a program to promote the application of big data analytics in the healthcare sector as an approach to dealing with problems encountered in such environments such as unnecessary admissions. Research indicates that the United States spends approximately $30 billion on unnecessary hospital admissions. In this regard, the aim of the Heritage Health competition was to seek innovators

who have the capability of developing predictive models, based on big data analytics. The developed models would play a role in predicting patients who are at a high risk of being admitted to hospital as well as the number of days that a patient. These models will enable physicians to establish ways to care for patients before their conditions worsen as well as to reduce the costs associated with unnecessary hospitalizations.[18]

Additionally, big data analytics has also been crucial in providing insights to care providers as well as patients in regards to compliance to the set standards of quality care treatment and predictive insights on the occurrence of adverse events. This application of big data analytics has been successful in institutions such as the Toronto hospital. The hospital management has established a system where they track the health condition of the patients using digital monitors that capture a wide variety of information such as the heartbeat rate, breathing patterns, blood pressure, and temperatures among other vital body indicators.[18] The information collected from these monitors has supported the development of predictive models that enable the physicians to foretell the likelihood of a patient's condition to deteriorate or the probability to suffer from an infection.

Big data analytics is also playing a significant role in reducing the rate of hospital readmissions. Statistics indicate that approximately 30% of all healthcare costs are as a result of hospitalization while more than a fifth of all case of hospital readmission occurs within 30 days after patients have been discharged.[18] These cases of readmission have led to a significant increase in the amount of costs arising from hospitalization. However, programs such as Medicare have taken advantage of the availability of hospital data to put in place measure to curb the rates of readmission. For instance, Medicare has put in place penalties against medical care institutions and providers that record a high rate of readmission.[20] On the other hand, some hospitals have developed systems that are based on big analytics to aid them in identifying patients who are at a high risk of getting readmitted.

Big data analytics has also played a role in enabling patients and consumers of health information to easily access information on healthcare conditions and self-management techniques and recommendations. The accessibility to this information has been enhanced by the availability of the internet and development of mobile search engines. However, this application of big data in healthcare may have negative legal implications given the fact that providers of this kind of data may not always provide evidence-based information and the treatment methods they provide to patients may not always be effective or correct.

The government has also been acting as a catalyst for increasing the rate of adoption of big data and big data analytics in the healthcare sector. Some of the initiatives that have been pioneered by governments include the General Electric Head Health Challenge whose aim was to promote the application of big data and big data analytics to advance research as well as diagnostic procedures of brain injuries.[3] This initiative was classified into two components. The first component was designed to promote research and innovation on this subject. This component aimed at seeking out individuals that would develop big data-based systems that would enhance the diagnosis procedures of mild head injuries as well as brain damage protection measures. The component also aimed at integrating bid data analytics-based measures towards reducing the occurrence and impact of head injuries in sports. The second component of the initiative was designed to enhance research on brain biomarkers that would enhance the process of Magnetic Resonance Imaging hence mitigating the occurrence and impact of brain injuries.[3]

The Indian government has also promoted big data analytics-based initiatives in the health sector by establishing and e-Intensive Care Unit (eICU) where patients can receive around the clock surveillance from teams of healthcare professionals who are located in a centralized unit.[21] This type of surveillance is made possible by the availability of remote sensors, imaging and video technologies that make it possible for the healthcare professionals to continuously monitor their patients remotely hence making it possible to act fast in the occurrence of life-threatening changes in their vitals. The eICU also features a clinical decision support system that allows healthcare professionals to make decisions.[21 easily]

A study conducted in regards to the eICU system using a representative sample of approximately 120,000 patients revealed that patients in the eICU were 26% more likely to survive compared to patients in the normal ICU environment.[3] The Indian government also supported a similar project dubbed The IntelliSpace Consultative Critical Care (ICCC) with the aim of attempting to address the challenge of lack of sufficient skilled ICU medical care personnel in the country mainly in the small towns and the rural areas.[3] This solution connected various hospitals in rural areas in a framework that would enable the available qualified and skilled physicians to treat patients remotely with the help of data and real-time images of the patients which were generated through video technology and monitoring sensors attached to the patients

International Journal of Engineering Applied Sciences and Technology, 2019
Vol. 4, Issue 5, ISSN No. 2455-2143, Pages 64-73
Published Online September 2019 in IJEAST (http://www.ijeast.com)

VII. THE ROLE OF BIG DATA AND BIG DATA ANALYTICS IN PROMOTING TECHNOLOGICAL INNOVATIONS IN THE HEALTHCARE SECTOR

The era of big data and big data analytics has also led to an increase in the number of healthcare innovators who are seeking to develop various technologies to improve health outcomes and the healthcare delivery process. An ideal example of a technological innovation that has resulted as a result of the big data revolution is y the GPS-enabled tracker which was manufactured by Asthmapolis.[3] This tracker was designed to track the inhaler usage patterns of individuals suffering from asthma. The data collected from the GPS trackers are stored in a central database. The accumulated patterns assist medical care providers in analysing the catalysts of this condition for certain individuals based on their inhaler usage patterns and locations. Information obtained from this technology plays a significant role in developing specialized or personalized treatments for individuals.[3]

The Ginger.io application is an example of another application that has been developed as a result of the big data revolution.[3] The developers of this platform seek the patients' consent to track them through their smartphones. This application allows the developer to track a patient's voice call patterns, messaging information, location, and movements, among others. Additionally, the developers can also send surveys on patients' devices and request them to respond. The data and information that is obtained from this application are compared from the patients' behavioral data, thus enabling the developers to make inferences about a patient's health condition. For instance, a change in a particular patient's communication or sleep patterns may be a sign that he or she is depressed or experiencing anxiety depending on the nature of the change. Additionally, reduced movements may be an indication that a certain patient is experiencing physical challenges. Such insights may be used to take medical action before an escalation of the health condition.

The mHealth platforms is another example of a technological innovation that leverages on the availability of big data to aid patients with chronic conditions through the provision of education and promotion of adherence to treatment.[3] This application obtains information and data on chronic conditions medication and their side effects as well as contacts of individuals who are under these medications. Such data is obtained from government websites such as clinicaltrial.gov and payors as well as providers. The developers then utilize this information to send out reminders as well as educative and targeted information to the identified individuals. Another example of innovation that leverages on big data technology was Optum's disease registry that utilized data from various sources to understand different diseases and their impacts on the patients.

VIII. THE PROCESS OF INTEGRATING BIG DATA AND BIG DATA ANALYTICS IN THE HEALTHCARE SECTOR

Despite the fact that big data and big data analytics have a significant potential to revolutionize the healthcare sector, healthcare organizations and providers need to understand that the process of integrating this new advancement into their operations. There are five major steps in the utilization of big data in the healthcare sector. These phases include data acquisition, data storage, data management, data analytics, and data visualization.[23]

a) Data Acquisition

The first phase in the process of implementing the use of big data analytics in the healthcare sector is acquiring the required data from the wide variety of available data. Big data is available in different forms, which include structured, unstructured, or semi-structured. Moreover, this data may be obtained from primary sources or secondary sources. The primary sources include electronic health record (EHR) systems, and medical decision support systems. On the other hand, the secondary sources include; insurance companies, government sources, and pharmacies among others.[23]

Nevertheless, one of the most common sources of big data in healthcare includes EHR systems. EHR systems have contributed significantly towards digitization of EHR records. EHR records may include prescription and medication history, patient history, laboratory reports, information from monitoring devices, and sensors, among others. This information may be used in analytics hence aiding in the development of personalized care and provision of crucial medical insights. Image processing technologies may also provide an important source of big data. Such technologies include Magnetic Resonance Imagery, X-rays, Computed tomography, ultrasounds, and mammography fluoroscopy, among others.

Healthcare organizations may also obtain big data from the internet and social media sites. Some of the crucial sources of this kind of information may include Facebook, Twitter, Instagram, and LinkedIn, among others. Moreover, some websites also act as an important source of big data for healthcare organizations. Some of the popular websites include 23andMe and uBiome. Information obtained from these sources is mainly used to track the spread or

outbreak of diseases. Moreover, healthcare organizations may also collect data from smartphones and other devices. These devices may contain applications that track an individuals steps, the number of calories burned each day, moods, and distance covered among others. Moreover, there are mobile phone applications for diabetic patients. The amount of health-related information generated by these devices is huge enough to be considered as big data.

b) Data Storage

It is also important for organizations to develop data storage mechanisms before they adopt the use of big data and big data analytics. The size of big data necessitates healthcare organizations to have sufficient space to store the information. Based on the size of big data, the most feasible storage option is the cloud.[24] Cloud storage not only provides sufficient storage but also ensure that the stored information can be accessed easily and quickly. Moreover, cloud storage also makes it possible to conduct computing procedures on the data. Using the cloud to store big data has a wide variety of advantages with the greatest one being the efficiency and cost-effectiveness of this form of storage. Additionally, there are numerous providers of cloud storage services. In this regard, organizations should ensure that they carefully select a cloud service provider who can be entrusted to protect the security and the privacy of the stored information. Some of the popular cloud service providers include Goggle, IBM, and Amazon.

c) Data Management

Healthcare organizations also need to understand the importance of efficient data management practices before they can adopt the use of big data analytics to aid in decision making and quality and healthcare quality improvement. The process of managing data entails data organization, data cleaning, data mining, and data governance. Data management also entails checking for missing data as well as verifying that all data is available. At this phase, data retrieval entails activities such as extracting valuable or required data from a huge database. On the hand, data governance entails all activities that are aimed at ensuring the security, accessibility, and usability of the stored information. This stage also entails maintaining the privacy and confidentiality of the stored information as well as adherence to the privacy and confidentiality laws and acts. There are multiple acts that regulate and influence the process of data governance. These Acts include HITECH, HDI, FOIA, GINA, and HIPPA.[8] The abbreviation HITECH stands for Health Information Technology for Economic and Clinical Health. This act stipulates the importance of acceptable use of health data as well as information. Its sets out the technological standards for enhancing the security and privacy of health data and encourages the implementation of electronic health records.

On the other hand, the Health Insurance Portability and Accountability Act also lays out the requirements for enhancing security, privacy, and confidentiality of stored health records. This act also establishes the national ethical standards for handling EHR records and other forms of communications and interactions within the healthcare setting. Despite the fact that the HIPPA acts mandates healthcare organization to protect their stored data, it also allows the transfers of data and patient records for various purpose as long as the right, and legal procedures are followed. The Department of Health and Human Services (HHS) has an initiative dubbed the Health Data Initiative (HDI) that seeks to reinforce the available health data security enhancement measures by allowing patients to rate the security of healthcare organizations based on their experiences

Furthermore, Genetic Information Nondisclosure Act of 2008 (GINA) forbids the use of or disclosure of genetic information for purposes that might disadvantage a patient such as underwriting or determination of premiums or the suitability for health insurance coverage. Lastly, the Freedom of Information Act (FOIA) also seeks ways of enhancing the security and privacy of stored data about various individuals.

d) Data Analytics

As mentioned before, data analytics is the process through which big raw data is transformed into relevant information that can guide decision making or provide important insights. The various approaches to big data analytics in the healthcare sector include Descriptive Analytics, Diagnostic Analytics, Predictive Analytics, and Prescriptive Analytics.[25] The Descriptive analytics method is based on historical data and examines the impact of a specific parameter in the system. Similarly, Diagnostic Analytics also relies on historical records to determine the cause of a certain observed medical problem. Additionally, the Predictive Analytics approach utilizes both historical and real-time data to make predictions about the future. The approach to data analytics is significantly probabilistic. Lastly, Prescriptive Analytics also utilizes historical records to make predictions about the multiple possible outcomes hence guiding in the process of decision making.[25]

e) Data Visualization

The last phase in the integration of big data analytics in the healthcare sector is a data visualization. Data visualization entails presenting the results of the different analytics approach in a manner that can be understood such a graphical or pictorial representation.[8] This stage makes it possible for

stakeholder in the healthcare environment to understand patterns and trends and to make informed decisions based on the observations.

IX. DATA VISUALIZATION AND ANALYTIC TOOLS

The are various data visualization and analytics tools that possess different capabilities. For instance, the Advanced Data Visualization tool has the capability to analyze data and present the results in graphical forms. This tool has can also scale its graphical representation to include millions of data points, and it also has the capability to handle different data types. Additionally, Presto is a different analytical tool that has the ability to analyze and present in visual forms huge sets of data that are collected each day by a certain system. Hive is a similar analytical tool to Presto. However, this tool is not as fast as Presto. Nevertheless, Hive is an efficient tool when excel performing Microsoft Excel functions.[26]. Additionally, Vertical is a cheaper alternative of the Hive application, given the fact that it has lesser architecture compared to vertical. However, Vertical has the aspect of scalability.

Moreover, Online Analytics Processing (OLAP) is another data analytics tool that has the capability to perform statistical calculations and reports, quality control, as well as check the integrity of the data. This system is ideal in improving the process of healthcare delivery and decision making, given the fact that it enhances tracking of medical records and past diagnoses. OLAP has similar attributes to Online Transaction Processing (OLTP). However, OLTP is mainly designated to process information related to operations within the healthcare setting, such as registration of patients. Furthermore, the Hadoop Distributed File System (HDFS) has the capability of dividing large data files into smaller ones and to redistribute these files across a range of different systems. HDFS has similar attributes to Casandra File System (CFS).[26]

Other tools include R; This programming tool enables graphical display and manipulation of data. Cytoscape is a similar data visualization tool that has advanced data modeling capabilities. Additionally, Graphviz is another application that is best suited for analytical results in both simple and complex graphical forms. Lastly, IBM cloud service providers also have a data visualization tool referred to as IBM Watson Analysis that is ideal when dealing with data stored in the IBM cloud storage. Other tools include Nodebox, Flot, FF Chartwell, Raphael, and Cross filter, Goggle Charts, MapReducing System, Complex Event Processing (CEP), Text Mining tools, Mahout and Nephi among others.[26]

X. CHALLENGES OF BIG DATA AND BIG DATA ANALYTICS IN THE HEALTHCARE SECTOR

Despite the various advantages and positive impacts that big data and big data analytics have on the healthcare sector, there are also various challenges associated with the adoption of this concept.[27] The major challenges associated with big data are mainly associated with data collection, usage, protection, and sharing.

One of the major challenges associated with big data is in regards to security.[27] As mentioned before, organizations need to have a big data storage plan before adopting big data analytics. The nature of healthcare big data is such that any leakage of information to the unintended parties may have serious and adverse legal implications to the involved healthcare organizations. Cloud storage, which is the ideal form of storage of big data is prone to various security challenges such as hacking, malware attacks, and phishing. This threat may cause the information to leak to malicious parties. In this regard, healthcare organizations need to put in place sufficient measures to safeguard their healthcare records. Moreover, cloud storage may be an expensive alternative storage option for small and medium healthcare organization which consequently limits their ability to leap from the benefits of big data analytics.[28] Moreover, the heterogeneous nature of big data may be strenuous for organizations to handle. Gathering the relevant as well as the required information from the huge pool of data may be a laborious and expensive process given the fact that it requires skilled personnel who possess knowledge about how to handle the various analytical and data mining tools.[28]

Moreover, training the hospital personnel on how to comfortably handle data analytics systems is an expensive and time-consuming endeavor.[29] Setting up of big data analytics systems may also be expensive, mainly for small scale organizations. In this regard, most organizations prefer to maintain their traditional systems. Moreover, the generational diversity that exists in the healthcare sector is also a significant challenge towards the adoption of big data and big data analytics.[30] The older generation of healthcare stakeholders may be opposed to the adoption of new systems of care delivery given the fact that most of them prefer traditional methods of care delivery as well as maintenance of hospital culture. Therefore, organizations that purpose to adopt the utilization of big data and big data analytics in their operations need to consider and take measures to mitigate these challenges.

XI. RESULTS AND DISCUSSIONS

The idea of big data in the healthcare industry is relatively new compared to other industries. Nevertheless, this concept is rapidly gaining popularity as most healthcare organizations and providers are gradually learning and appreciating the capability of big data in improving care delivery and patient outcomes. The size of medical data has been on the rise in recent years. The size of medical data in 2012 was approximately 500 petabytes. This size is expected to increase to 25, 000 before 2020. The data is also expected to increase in variety and complexity. The increase in size, variety, and complexity of medical data has led most hospitals to begin making decisions based on this data. Current trends in digitization, which have been fueled by the increased in demand for quality healthcare services and the need to improve the efficiency of the healthcare delivery process have led the availability of data in soft-copy hence giving rise to the era of big data in the healthcare industry.

There are various sources of big data in the healthcare setting these include notes from physicians, medical prescription records, imaging technologies, laboratory data, data from the pharmacy, insurance data, and other administrative and medical process records. Additionally, big data in the healthcare environment may be generated from Electronic Health Records, medical sensors data, customers feedbacks and posts on social media sites such as Facebook, Twitter, Instagram, websites activities and news feeds among others. Moreover, there are multiple characteristics associated with big data in the healthcare industry. They include velocity, volume, variety, variability, veracity, and value

Furthermore, there are promising prospects for big data and big data analytics in the health care sector. Big data and big data analytics can be used in supporting clinical decision making. Improved decision-making leads to better patient relationships as well as outcomes. Improved and informed decision making also helps to reduce the chances for error. Moreover, Big data analytics has the capability of improving the provision of personalized care to patients by providing healthcare providers with real-time information about patients hence allowing them to understand a patient's condition and to engage in best practices when dealing with the identified problems. Big data analytics can also be used to visualize data on the spread and outbreak of disease and to predict future patterns in regards to such occurrences. Big data analytics also has the capability to improve clinical operations and care delivery systems.

Furthermore, research indicates that there are various cases of successful application of big data analytics in the healthcare sector in different regions and platforms around the world. One such case is the application of big data, and big data analytics to monitor the outbreak and spread of diseases. Google has successfully been able to predict disease outbreaks by analyzing search engine queries in different locations and their frequency. Healthcare organizations are applying this capability of big data analytics to analyze the contents of the huge database of medical research and literature hence allowing physicians to be able to make more informed decisions. Big data analytics can also help individuals to understand about their own conditions by presenting information in a manner that is easy to understand hence allowing such individuals to participate in decision making in regards to their health conditions.

Furthermore, institutions such as Heritage Health have initiated a program to promote the application of big data analytics in the healthcare sector as an approach to dealing with problems encountered in such environments such as unnecessary admissions. Big data analytics is also playing a significant role in reducing the rate of hospital readmissions. For instance, Medicare has put in place penalties against medical care institutions and providers that record a high rate of readmission. On the other hand, some hospitals have developed systems that are based on big analytics to aid them in identifying patients who are at a high risk of getting readmitted

The application of big data analytics has been successful in institutions such as the Toronto hospital. The hospital management has established a system where they track the health condition of the patients using digital monitors that capture a wide variety of information such as the heartbeat rate, breathing patterns, blood pressure, and temperatures among other vital body indicators. The information collected from these monitors has supported the development of predictive models that enable the physicians to foretell the likelihood of a patient's condition to deteriorate or the probability to suffer from an infection.

The government has also been acting as a catalyst towards increasing the rate of adoption of big data and big data analytics in the healthcare sector by sponsoring initiatives designed to promote usage of big data analytics in the healthcare sector. The era of big data and big data analytics has also led to an increase in the number of healthcare innovators who are seeking to develop various technologies to improve health outcomes and the healthcare delivery process.

Despite the fact that big data and big data analytics have a significant potential to revolutionize

the healthcare sector, healthcare organizations and providers need to understand the process of integrating this new advancement in to their operations by following the five major phases which include data acquisition, data storage, data management, data analytics and data visualization. Moreover, there are various tools and systems that organizations can utilize to aid in data analytics and visualizations.

Despite the various advantages and positive impacts that big data and big data analytics has on the healthcare sector, there are also various challenges associated with the adoption of this concept. The major challenges associated with big data are mainly associated with data collection, usage, protection, sharing, as well as the costs associated with big data and big data analytics. Therefore, organizations that purpose to adopt the utilization of big data and big data analytics in their operations need to consider and take measures to mitigate these challenges.

XII. CONCLUSION AND RECOMMENDATIONS

Despite the slow adoption of big data and big data analytics in the healthcare sector, this concept has slowly been gaining momentum. More healthcare organizations are slowly realizing the importance of big data in improving healthcare outcomes and process. Based on the literature review, big data and big data analytics has the capability to improve decision making in the healthcare sector, predict disease outbreaks as well as the trends and patterns of the spread of such diseases, predict occurrence of medical phenomenon's such as hospital readmission, reoccurrence of diseases, and risk of infection among others. Moreover, big data analytics has the capability to help healthcare organizations to streamline processes within the healthcare setting. However, the process of integrating big data analytics in the healthcare setting follows distinct phases. Healthcare organizations also have to consider the challenges associated with adopting big data analytics. Nevertheless, based on the literature, big data analytics has the capability to improve delivery systems and outcomes within the healthcare sector.

XIII. REFERENCES

[1] Monino, J. (2016). Data Value, Big Data Analytics, and Decision-Making. *Journal of the Knowledge Economy*. doi:10.1007/s13132-016-0396-2

[2] Panda, D. K. (2015). Accelerating Big Data Processing on Modern Clusters. Proceedings of the 1st Workshop on Performance Analysis of Big Data Systems - PABS '15. doi:10.1145/2694730.2694733.

[3] Dang, A., & Mendon, S. (2015). The Value of Big Data in Clinical Decision Making. *International Journal of Computer Science and Information Technologies, 6*(4), 3830-3835.

[4] Huang, T., Lan, L., Fang, X., An, P., Min, J., & Wang, F. (2015). Promises and Challenges of Big Data Computing in Health Sciences. *Big Data Research*, *2*(1), 2-11. doi:10.1016/j.bdr.2015.02.002

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

[6] Fan, J., Han, F., & Liu, H. (2014). Challenges of Big Data Analysis. *National science review*, *1*(2), 293–314. doi:10.1093/nsr/nwt032

[7] El Aboudi, N., & Benhlima, L. (2018). Big Data Management for Healthcare Systems: Architecture, Requirements, and Implementation. *Advances in bioinformatics*, *2018*, 4059018. doi:10.1155/2018/4059018

[8] SA, S., Rai, B. K., Meshram, A. M., & Gunasekeran, A. (2018). Big Data in Healthcare Management: A Review of Literature. *American Journal of Theoretical and Applied Business*, *4*(2), 57.

[9] El aboudi, N., & Benhlima, L. (2018). Big Data Management for Healthcare Systems: Architecture, Requirements, and Implementation. *Advances in Bioinformatics*, 2018, 1-10. doi:10.1155/2018/4059018

[10] Juddoo, S., George, C., Duquenoy, P., & Windridge, D. (2018). Data Governance in the Health Industry: Investigating Data Quality Dimensions within a Big Data Context. *Applied System Innovation*, *1*(4), 43. doi:10.3390/asi1040043

[11] Dagliati, A., Tibollo, V., Sacchi, L., Malovini, A., Limongelli, I., Gabetta, M., … Bellazzi, R. (2018). Big Data as a Driver for Clinical Decision Support Systems: A Learning Health Systems Perspective. Frontiers in Digital Humanities, 5. doi:10.3389/fdigh.2018.00008

[12] Bae, J. (2017). Shared decision making: relevant concepts and facilitating strategies. *Epidemiology and health*, *39*, e2017048. doi:10.4178/epih.e2017048

[13] Alotaibi, Y. K., & Federico, F. (2017). The impact of health information technology on

patient safety. *Saudi medical journal*, *38*(12), 1173–1180. doi:10.15537/smj.2017.12.20631

[14] Islam, M. S., Hasan, M. M., Wang, X., Germack, H. D., & Noor-E-Alam, M. (2018). A Systematic Review on Healthcare Analytics: Application and Theoretical Perspective of Data Mining. *Healthcare (Basel, Switzerland)*, *6*(2), 54. doi:10.3390/healthcare6020054

[15] Bansal, S., Chowell, G., Simonsen, L., Vespignani, A., & Viboud, C. (2016). Big Data for Infectious Disease Surveillance and Modeling. *The Journal of infectious diseases*, *214*(suppl_4), S375–S379. doi:10.1093/infdis/jiw400

[16] Palanisamy, V., & Thirunavukarasu, R. (2017). Implications of big data analytics in developing healthcare frameworks – A review. *Journal of King Saud University - Computer and Information Sciences*. doi:10.1016/j.jksuci.2017.12.007

[17] Verma, M., Kishore, K., Kumar, M., Sondh, A. R., Aggarwal, G., & Kathirvel, S. (2018). Google Search Trends Predicting Disease Outbreaks: An Analysis from India. *Healthcare informatics research*, *24*(4), 300–308. doi:10.4258/hir.2018.24.4.300.

[18] Priyanka, K., & Kulennavar, N. (2014). A Survey on Big Data Analytics in Health Care. *International Journal of Computer Science and Information Technologies*, *5*(4), 5865-5868.

[19] Zhu, R., Han, S., Su, Y., Zhang, C., Yu, Q., & Duan, Z. (2019). The application of big data and the development of nursing science: A discussion paper. *International Journal of Nursing Sciences*, *6*(2), 229-234. doi:10.1016/j.ijnss.2019.03.001

[20] Gupta, S., Zengul, F. D., Davlyatov, G. K., & Weech-Maldonado, R. (2019). Reduction in Hospitals' Readmission Rates: Role of Hospital-Based Skilled Nursing Facilities. *Inquiry: a journal of medical care organization, provision and financing*, *56*, 46958018817994. doi:10.1177/0046958018817994

[21] Scurlock, C., & Becker, C. (2016). Telemedicine for Trauma and Emergency: the eICU. Current Trauma Reports, 2(3), 132-137. doi:10.1007/s40719-016-0054-

[22] Chen, M., Mao, S., Zhang, Y., & Leung, V. C. (2014). Big Data Generation and Acquisition. Big Data, 19-32. doi:10.1007/978-3-319-06245-7_3

[23] Lyko K., Nitzschke M., Ngonga Ngomo AC. (2016) Big Data Acquisition. In: Cavanillas J., Curry E., Wahlster W. (eds) New Horizons for a Data-Driven Economy. Springer, Cham

[24] Strohbach M., Daubert J., Ravkin H., Lischka M. (2016) Big Data Storage. In: Cavanillas J., Curry E., Wahlster W. (eds) New Horizons for a Data-Driven Economy. Springer, Cham

[25] Vassakis, K., Petrakis, E., & Kopanakis, I. (2017). Big Data Analytics: Applications, Prospects and Challenges. Mobile Big Data, 3-20. doi:10.1007/978-3-319-67925-9_1

[26] Alkhatib, M. A., Talaei-Khoei, A., & Ghapanchi, A. H. (2015). *Analysis of Research in Healthcare Data Analytics*. Sydney: Australasian Conference on Information Systems.

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

[28] Ristevski, B., & Chen, M. (2018). Big Data Analytics in Medicine and Healthcare. *Journal of Integrative Bioinformatics*, *15*(3). doi:10.1515/jib-2017-0030

[29] Adibuzzaman, M., DeLaurentis, P., Hill, J., & Benneyworth, B. D. (2018). Big data in healthcare - the promises, challenges and opportunities from a research perspective: A case study with a model database. *AMIA ... Annual Symposium proceedings. AMIA Symposium*, *2017*, 384–392.

[30] Belle, A., Thiagarajan, R., Soroushmehr, S. M., Navidi, F., Beard, D. A., & Najarian, K. (2015). Big Data Analytics in Healthcare. *BioMed Research International*, *2015*, 1-16. doi:10.1155/2015/370194