International Journal of Innovative Approaches in Education
2018, Vol. 2 (1), 13-22
https://doi.org/10.29329/ijiape.2018.134.2

Original article

A Survey on Reducing Costs Vs Improving Outcomes and Opportunities in Health Care Field Using Big Data

S. P. Santhoshkumar [a, *], M. Praveen Kumar [a], S. Karthick [a] & S. Syed Shajahaan [a]

[a] Department of CSE, Rathinam Technical Campus, Coimbatore, Tamilnadu, India.

Abstract

In recent years, the total amount of medical data collected has grown exponentially, mainly because of the introduction of new medical technologies and the increased use of IT to manage medical health records of patients. These technologies have been introduced not only to provide higher quality of care but also to help solve some of the major healthcare problems in developed countries: the lack of productivity and efficiency, the increasing costs and the relatively low accessibility of the system. Big Data, or the capability to develop, analyze and visualize huge and complex data sets for various research interests and applications, has the promise to radically impact the speed and direction in tackling these problems. Health care field grows tremendously in last few decades. The health care field has generated huge amounts of data that has huge volume, enormous velocity and vast variety. Also it comes from a variety of new sources as hospitals are now tend to implemented electronic health record (EHR) systems. These sources have strained the existing capabilities of existing conventional relational database management systems. In such scenario, Big Data solutions offer to harness these massive, heterogeneous and complex data sets to obtain more meaningful and knowledgeable information. This paper basically studies the impact of implementing the big data solutions on the healthcare field, the potential opportunities, challenges and available platform and tools to implement Big data analytics in health care field.

Keywords: Electronic Health Record, Medical Health Records, Big Data, Health Care, Big Data Analytics.

Received: 15 March 2018 * **Accepted**: 19 March 2018 * **DOI**: https://doi.org/10.29329/ijiape.2018.134.2

* **Corresponding author:**

S. P. Santhoshkumar, Assist. Prof. Dr., Department of CSE, Rathinam Technical Campus, Coimbatore, Tamilnadu, India.
Email: spsanthoshkumar16@gmail.com

INTRODUCTION

The health care sector grows rapidly in last 30 years. The healthcare industry historically has generated large amounts of data, driven by record keeping, compliance & regulatory requirements and patient care. While most data is stored in hard copy form, the current trend is towards the rapid digitization of these large amounts of data. There are different types of data sources which generates these enormous amounts of data. Big data in healthcare refers to electronic health care records (EHR) that is quite large and complex that they are difficult to manage with traditional software and/or hardware. Also, they are not easily managed with traditional or common data management tools and methods. Using the technologies that able to deal with such "Big Data" will offer many potential opportunities to the healthcare sector.

An electronic health record (EHR), or electronic medical record (EMR), is the systematized collection of patient and population electronically-stored health information in a digital format. These records can be shared across different health care settings. Records are shared through network-connected, enterprise-wide information systems or other information networks and exchanges. EHRs may include a range of data, including demographics, medical history, medication and allergies, immunization status, laboratory test results, radiology images, vital signs, personal statistics like age and weight, and billing information.

EHR systems are designed to store data accurately and to capture the state of a patient across time. It eliminates the need to track down a patient's previous paper medical records and assists in ensuring data is accurate and legible. It can reduce risk of data replication as there is only one modifiable file, which means the file is more likely up to date, and decreases risk of lost paperwork. Due to the digital information being searchable and in a single file, EMRs are more effective when extracting medical data for the examination of possible trends and long term changes in a patient. Population-based studies of medical records may also be facilitated by the widespread adoption of EHRs and EMRs.

The Motivation for Big Data

Health care costs are driving the demand for big-data driven Healthcare applications. U.S. health care spending has outpaced GDP growth for the past several decades and exceeds spending in any other developed country. Despite being more expensive, according to the Organisation for Economic Co-operation and Development (OECD), the US Health System ranks last among eleven countries on measures of access, equity, quality, efficiency, and healthy lives. Standards and incentives for the digitizing and sharing of healthcare data along with improvements and decreasing costs in storage and parallel processing on commodity hardware, are causing a big data revolution in health care with the goal of better care at lower cost.

Value Based Care

A goal of the Affordable Care Act is to improve health care through the meaningful use of health information technology in order to:

a. Improve healthcare quality and coordination so that outcomes are consistent with current professional knowledge

b. Reduce healthcare costs, reduce avoidable overuse

c. Provide support for reformed payment structures

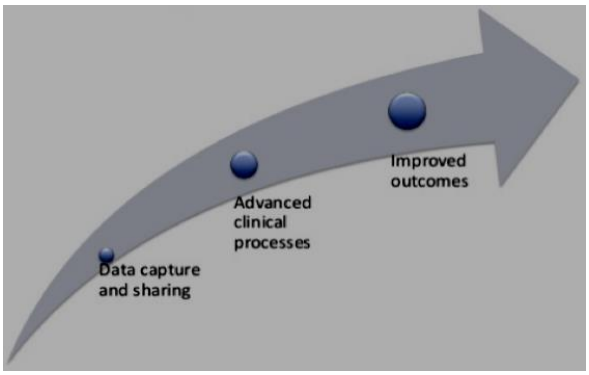

Figure 1: Healthcare Quality Outcomes

Health Insurance companies, Medicare and Medicaid are shifting from fee-for-service compensation to value based data driven incentives that reward high quality, cost effective patient care and demonstrate meaningful use of electronic health records.

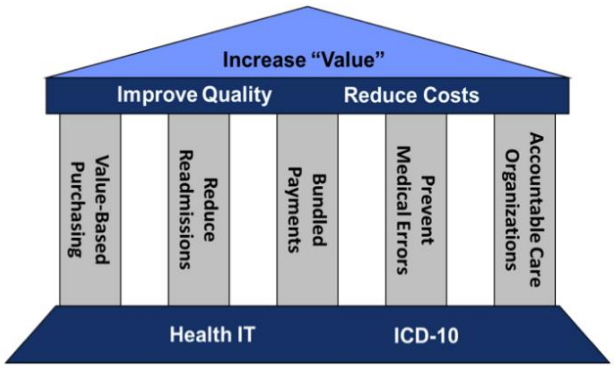

Figure 2: Increase Value and Reduce Cost

Health Care Data

Unstructured data forms about 80% of information in the healthcare industry and is growing exponentially. Getting access to this unstructured data—such as output from medical devices, doctor's

notes, lab results, imaging reports, medical correspondence, clinical data, and financial data—is an invaluable resource for improving patient care and increasing efficiency.

Examples of healthcare data sources that will benefit from big data and analytics:

- *Claims:* are the documents providers submit to insurance companies to get paid. A key component of the Health Insurance Portability and Accountability Act (HIPAA) is the establishment of national standards for electronic healthcare transactions in order to improve efficiency by encouraging the widespread use of Electronic Document Interchange (EDI) between healthcare providers and insurance companies. Claim transactions include International Classification of Diseases (ICD) diagnostic codes, medications, dates, provider IDs, the cost.
- *Electronic Health/Medical Record data (EHR or EMR):* Medicare and Medicaid EHR incentive programs were established to encourage professionals and hospitals to adopt and demonstrate meaningful use of certified EHR technology. EHRs facilitate a comprehensive sharing of data with other providers and medical applications. EHRs contain the data from the delivery of healthcare which includes diagnosis, treatment, prescriptions, lab tests, and radiology. Health Level Seven International (HL7) provides standards for the exchange, integration, sharing, and retrieval of electronic health record data.
- *Pharmaceutical R&D:* Clinical Trials Data, Genomic Data.
- Patient behavior and sentiment data.
- Medical Device Data: Patient sensor data from the home or hospital.

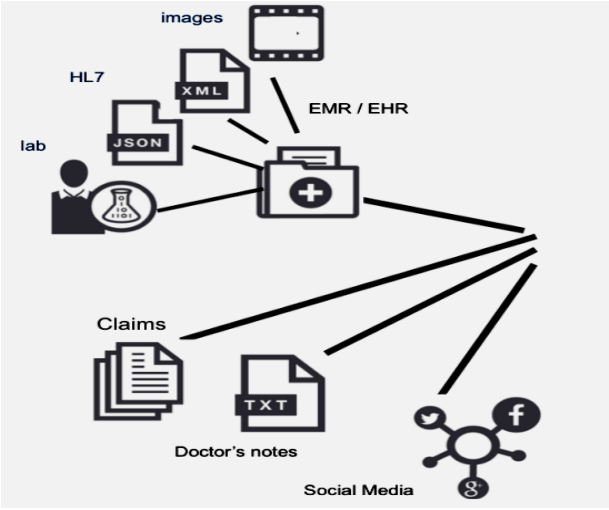

Figure 3: Healthcare Data

Big Data Trends in Healthcare

There is a move toward evidence-based medicine, which involves making use of all clinical data available and factoring that into clinical and advanced analytics. Capturing and bringing all of the information about a patient together gives a more complete view for insight into care coordination and outcomes-based reimbursement, population health management, and patient engagement and outreach.

Reducing Fraud Waste and Abuse with Big Data Analytics

The cost of fraud, waste and abuse in the healthcare industry is a key contributor to spiraling health care costs in the United States, but big data analytics can be a game changer for health care fraud. The Centers for Medicare and Medicaid Services prevented more than $210.7 million in healthcare fraud in one year using predictive analytics. United Healthcare transitioned to a predictive modeling environment based on a Hadoop big data platform, in order to identify inaccurate claims in a systematic, repeatable way and generated a 2200% return on their big data/advanced technology.

The key to identifying fraud is the ability to store and go back in history to analyze large unstructured datasets of historical claims and to use machine-learning algorithms to detect anomalies and patterns.

Healthcare organizations can analyze patient records and billing to detect anomalies such as a hospital's overutilization of services in short time periods, patients receiving healthcare services from different hospitals in different locations simultaneously, or identical prescriptions for the same patient filled in multiple locations.

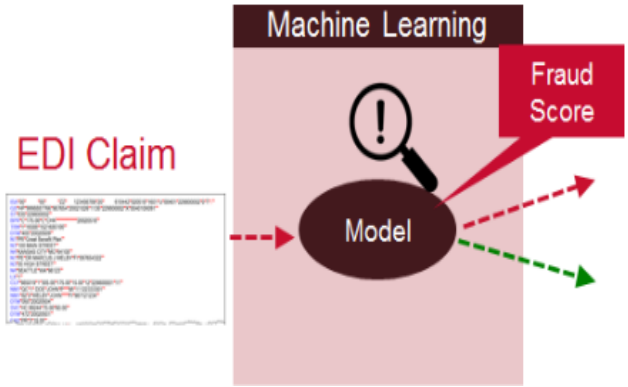

Figure 4: Reducing Fraud using Machine Learning

The Centers for Medicare and Medicaid Services uses predictive analytics to assign risk scores to specific claims and providers, to identify billing patterns, and claim aberrancies difficult to detect by previous methods. Rules-based models flag certain charges automatically. Anomaly models raise suspicion based on factors that seem improbable. Predictive models compare charges against a fraud

profile and raise suspicion. Graph models raise suspicion based on the relations of a provider; fraudulent billers are often organized as tight networks.

Predictive Analytics to Improve Outcomes

Initiatives such as meaningful use are accelerating the adoption of Electronic Health Records and the volume and detail of patient information is growing rapidly. Being able to combine and analyze a variety of structured and unstructured data across multiple data sources, aids in the accuracy of diagnosing patient conditions, matching treatments with outcomes, and predicting patients at risk for disease or readmission.

Predictive modeling over data derived from EHRs is being used for early diagnosis and is reducing mortality rates from problems such as congestive heart failure and sepsis. Congestive Heart Failure (CHF) accounts for the most health care spending. The earlier it is diagnosed the better it can be treated avoiding expensive complications, but early manifestations can be easily missed by physicians. A machine learning example from Georgia Tech demonstrated that machine-learning algorithms could look at many more factors in patients' charts than doctors, and by adding additional features there was a substantial increase in the ability of the model to distinguish people who have CHF from people who don't.

Predictive modeling and machine learning on large sample sizes, with more patient data, can uncover nuances and patterns that couldn't be previously uncovered. Optum Labs has collected EHRs of over 30 million patients to create a database for predictive analytics tools that will help doctors make Big Data-informed decisions to improve patients' treatment.

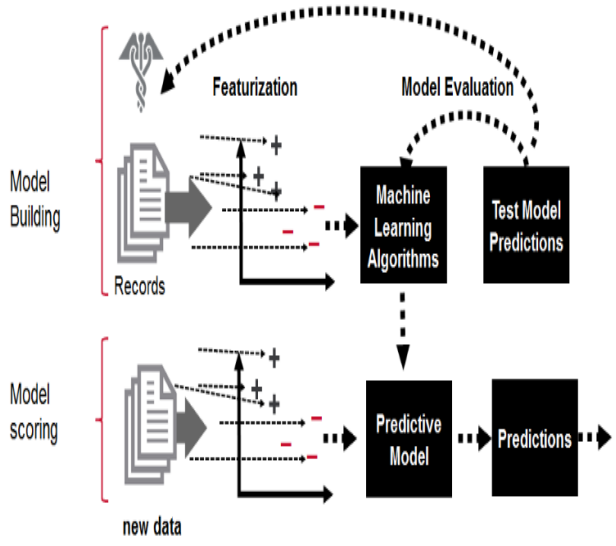

Figure 5: Predictive Analytics to Improve Outcomes

Real-Time Monitoring of Patients

Healthcare facilities are looking to provide more proactive care to their patients by constantly monitoring patient vital signs. The data from these various monitors can be analyzed in real time and send alerts to care providers so they know instantly about changes in a patient's condition. Processing real-time events with machine learning algorithms can provide physicians' insights to make lifesaving decisions and allow for effective interventions.

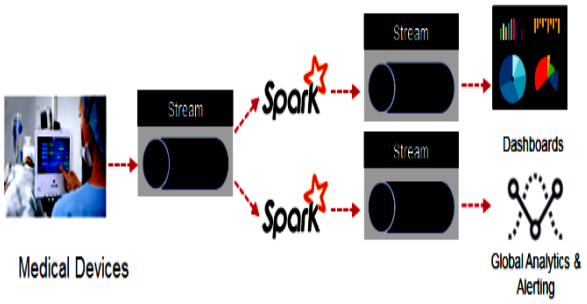

Figure 6: Real-Time Monitoring of Patients

Health Care and Big Data

An information and communications technology (ICT) is playing a vital role in improving health care for individuals and communities. It helps to improve health system efficiencies and prevent medical errors. With an invent of new and efficient mechanisms for storing and accessing information, ICT helps to serve a society in a better way. ICT powered health mechanisms are often known as eHealth.

One of the characteristic that health care sector possesses is its data richness. With the development in diagnostic and treatment, health care sector evolved so quickly in last few decades. There are many sources in this sector from where the data is generated. These data is

Undoubtedly in the form of Big Data. The data came from many sources and categorized as follows:

- Web and social media data: Data captured from Facebook, Twitter, LinkedIn, blogs, and the like. It can also include health plan websites, Smartphone apps etc.
- Machine-to-machine (M2M) device generated data: readings from remote sensors, meters, and other devices [7].
- Biometric data: Data may in form of retinal scans, x-ray images, finger prints, genetics, handwriting, other medical images, blood pressure and other similar types of data.
- Human-generated data: In the form of unstructured and semi-structured data. Some of the examples are EMRs, Doctor's notes and paper documents.

- Genomic Data: data in the form of DNA sequence [3].

Opportunities of Big Data in Healthcare

This section discusses the various opportunities of Big data in health care.

Decreasing Healthcare Costs to Get Financial Profit

Big data can help decrease the cost of providing medical treatment in many ways. Moreover, analysis on data gives insight to health care providers to determine populations at risk for illness. By doing so, proactive steps can be taken initially. Data and its analytics are easier than ever to share. Big data can more accurately pinpoint where education and prevention is needed to produce healthier populations at lower costs. Treatment is more evidence based using Big Data analytics [5].

Promotes Research and Innovation

By analytics on data, the current state of health of patients provides insight to them to take more ownership of their healthcare. The information sharing mechanism increases productivity and reducing overlapping of data. By thus, it is enhancing the coordination of care. [5].

Personalized Medicine

In past few years, it is possible to predict the lifestyle diseases through genetic blue prints. Big data will further personalize medicine by determining the tests and treatments needed for each patient. The provision of earlier treatment can reduce the health costs and can eliminate the risk of chronic diseases [5].

Strengthen the Preventive Care

Prevention is always better than cure. Following this thumb of rule, with the advent of Big Data analytics, it is easy to capture, analyze and compare patient symptoms earlier to offer a preventive care in a better way.

Virtual Care and Wearable Health Care

Technologies Technology is helping providers make virtual care initiatives that increase quality of care and provide patients with more access [4].

Health Trend Analysis

By using different analytical approaches including data mining and text mining techniques, health trend analysis and comprehensive patient management is more easy using Big Data Analytics [8]

Identification and Tracking of Patients

The identification and tracking of patients with type 2 diabetes is discussed in recent article [7]. The author suggests to use a two-step process to identify subsets of patients that have similar clinical

indications and care patterns. In a first step, patients are divided into groups based on the primary diagnosis. Then after, a statistical clustering method is applied to further divide the subsets. This method uses readily available administrative datasets. Also, patients must be tracked longitudinally to determine the patterns for treatment. Therefore, the method is applicable in scenarios where patient data is available over time and across providers.

Studying Drug Efficacy

Electronic health record (EHR) data may also be used to study drug efficacy. Researchers at the University Of Pennsylvania School Of Medicine compared the results of randomized controlled trials versus using an EMR to compare cardiovascular outcomes. It has been observed that the cost of randomized controlled trials is much higher than the cost of using readily available EHR data to compare treatment modalities.

Conclusion

We may consider Big data as a latest evolution in the field of decision support data management systems. On the other side, the digitalization in health care sector is in peak. As we discussed in the paper, there are several opportunities for big data in health care sector. Meanwhile, the technological advancement is rapidly going on towards the implementation of big data analytics. In near future, there will be widespread implementation of big data analytics across the healthcare organization and the healthcare industry. The Big data solutions could definitely save millions of life and improve patient services.

REFERENCES

Sanskruti Patel and Atul Patel, "A big data revolution in health care sector: opportunities, challenges and technological advancements", International Journal of Information Sciences and Techniques (IJIST) Vol.6, No.1/2, March 2016

Big data analytics definition, retrieved from http://searchbusinessanalytics.techtarget.com/definition/big-data-analytics.

Big Data Analytics for Healthcare, retrieved from https://www.siam.org/meetings/sdm13/sun.pdf.

Big Data in health care, retrieved from http://www.sas.com/en_us/insights/articles/big-data/big-datain-healthcare.html

International Journal of Information Sciences and Techniques (IJIST) Vol.6, No.1/2, March 2016

Big Data Offers Big Opportunities in Healthcare, retrieved from http://www.villanovau.com/resources/bi/big-data-healthcare-opportunities/#.VnfRArZ95kg

Big Data, retrieved from http://www.sas.com/en_us/insights/big-data/what-is-big-data.html on December 20, 2015

Bradley P. Implications of big data analytics on population health management. Big Data. 2013;1(3):152–159.

C. Chai, "Big Data in Healthcare", retrieved from http://ihealthtran.com/pdf/Big%20Data%20in%20Healthcare_Chai.pdf

Doug Laney, Application Delivery Strategies, retrieved from http://blogs.gartner.com /douglaney/files/2012/01/ad949-3D-Data-[11]. Management-Controlling-Data-Volume-Velocity-andVariety.pdf on December 20, 2015.

Hugh J. Watson, Tutorial: Big Data Analytics: Concepts, Technologies, and Applications, Communications of the Association for Information Systems, Volume 34, Article 65, pp. 1247-1268, April 2014.

Informatics or Analytics? Understanding Digital Health Use Cases, retrieved fromhttp://practicalanalytics.co/2013/07/15/informatics-or-analytics-understanding-healthcare-provideruse-cases/, HIMSS 2013.