An Empirical Study: Data Analytics Fosters Health Care Data Management And Its Outcomes

Saravanan C

Savita Sheelavant

Assistant Professor, Department of MCA,
R. V. College of Engineering, Bengaluru

Abstract: In Recent times, Healthcare industry has become one of the India's largest sectors in terms of profits and employment. Its environmental impact has become an important factor globally and is continuing to draw the attention of regulators. The Amount of health care data generated is increasing in alarming rate. According to the IDC Health insights, worldwide health care data is estimated to grow to 25,000 petabytes by 2020 a 50-fold increase within eight years, wherein it was 500 petabytes in 2012. According to FICCI -KPMG report, Indian health sector which was $73.92 billion in 2011 is expected to be $280 billion in size by 2020 with compound annual growth rate of 16 percent. The Indian healthcare workforce is expected to double to 7.4 million in 2022 from 3.6 million in 2013. This paper provides an insight about the importance of health care data, sources of health care data generation and formats, challenges and trends of handling it. This paper also focuses on how big data analytics can help in improving the real-time monitoring of patients and outcomes of health care data.

Keywords: Health care, Big data, Big data analytics, Patient-Centric

I. INTRODUCTION

A health care system is the organization of people, institutions and resources that provide health care services to meet the health needs of target populations. It comprises of hospitals, medical devices, clinical trials, outsourcing, telemedicine, medical tourism, medical insurance and medical equipment. Hospitals can be classified into Government and private hospitals. Government hospitals include health care centers, district hospitals and general hospitals. Private hospitals include nursing homes, mid-tier and top tier private hospitals. Medical insurance consists of health insurance and medical reimbursement facility which includes an individual's hospitalization treatment expenses incurred due to sickness. Medical Equipment includes organizations mainly manufacturing medical devices and equipment like laboratory instruments, dental, surgical instruments, orthopedic instruments. Figure 1 represents the various components of health care system. Improvements in the health of citizens contribute to overall economic prosperity of the nation.
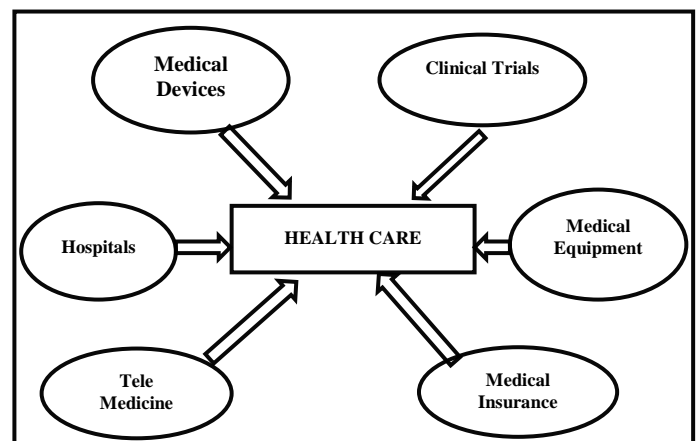

Figure 1: Health care system components

II. CHARACTERISTICS OF HEALTH CARE DATA

The health care industries have high amount of data set collections about diagnosis, patient details and medications [1]. The Healthcare Data collected is unique and complex in nature. It's difficult to Measure. Health care data is obtained from multiple locations and is available in different formats such as text, numeric, paper, digital, pictures, videos and multi medias. Health Care Data is being produced continuously and consumed in various formats, sizes from several different sources. The rate at which medical data collected in health care sector is virtually continuous. A health care Data sets generated in Health informatics applications are very complex for storing, processing and analysis. [2]. Large amount of health care information created in health care sector from electronic and smart devices would be equivalent to 500 billion file cabinets. Table 1 presents the various format of health care data available from different sources

The health care data is structured as well as unstructured. Structured data is consistent and exist in pre-defined fields within the record. It is computable, analyzable, comparable, quantifiable, and reportable [8]. Structured medical data in health care would be a temperature, weight, blood pressure reading, glucose level reading, red blood cell count, laboratory result value or patient demographic data. Example of structured health data can be relational database consisting of various tables storing the information of patients, drugs, treatment cost and so on. Unstructured data is the information that is challenging to organize and to be stored using traditional database. Example of unstructured medical data can be X ray images, Physicians notes, MRI, CT and other medical Correspondence as shown in Figure 2, 3 and 4. Another example of unstructured data in Health IT can be paragraph which represents the history of patient illness. "*The patient came in complaining of chest pain, shortness of breath, and lingering headaches…smokes 2 packs a day… family history of heart disease… has been experiencing similar symptoms for the past hours*".

Figure 2: Doctors Note sample
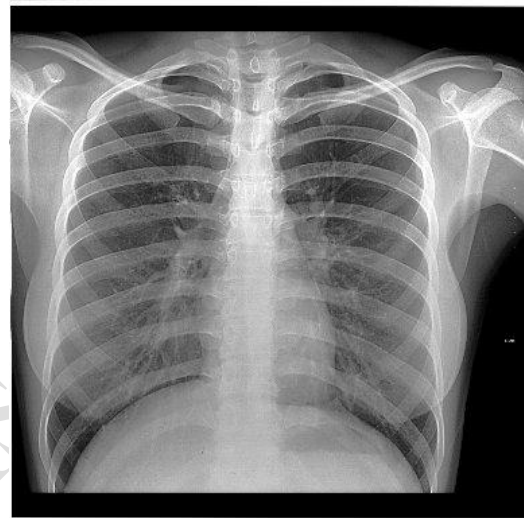
Figure 3: X-Ray sample

Figure 4: MRI sample

Varied data	Source
Administrative	*At patient level* - Identification, demographics, diagnosis *At provider level* -Procedures, costs, charges, bill payments
Clinical / Electronic Medical Records	Lab test results, Prescriptions, Pathology reports, Clinician notes
External Sources	Research, Pharmacy, Government, Population.
Medical Device Generated	Blood pressure, heart rate, glucose levels, weight.
Imaging & Diagnostics	X-ray, Computerized Tomography(CT) and Magnetic Resonance Imaging (MRI)
Customer Experience	Web, Call Center, Social Media Networks

Table 1: Sources of Health Care data from various Sources
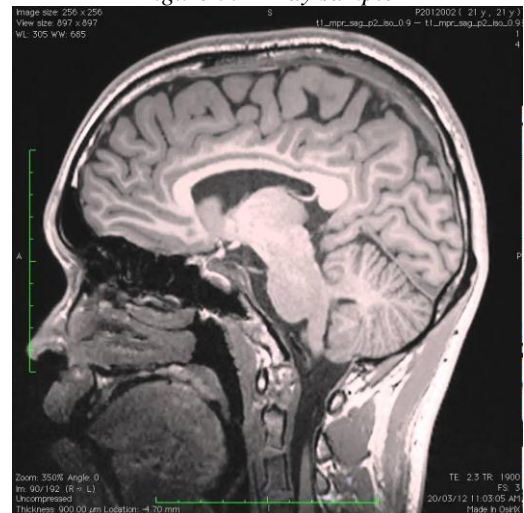
III. PURPOSE OF ANALYZING THE HEALTH CARE DATA

The healthcare data collected from various sources with different formats can be analyzed for identifying the patterns and trends, cost considerations, disease surveillance, improving outcome of patient heath, health management to larger population, providing clinical decision support, research, reduce risk and detection of frauds.

IV. ADVANTAGES OF BIGDATA ANALYTICS IN HEALTH CARE DOMAIN

- ✓ *CLINICAL OPERATIONS:* To decide on choosing best procedures for providing optimal clinical processes to diagnose and treat patients effectively.
- ✓ *EFFECTIVENESS IN RESEARCH:* Comparative effectiveness research is designed to inform health-care decisions by providing evidence on medical intervention at the population level, the effectiveness, benefits, and harms of different treatment options [3]. The evidence is generated from research studies that compare drugs, medical devices, tests, surgeries, or ways to deliver health.
- ✓ *USAGE AND ANALYSIS OF PUBLIC HEALTH DATA:* To analyze population wide effective and quality treatment, financial structure, identifying trends, outcomes, disease patterns, tracking disease outbreaks to influence patient care, drug development for more accurately targeted vaccines. For the analysis, the aggregated data collected form clinical, financial, administrative and other related heath data of patient's specific disease diagnosis and procedures.
- ✓ *FRAUD DETECTION:* Healthcare fraud is very serious financial problem faced by healthcare systems. It represents a serious problem on the effectiveness of providing better healthcare to needy individuals. Data analysis provides a well effective way to be more proactive in fraud detection and prevention in health care. Some of the fraud schemes encountered and prevented are listed below [4].
  - Generating Bills for unnecessary medical tests
  - Identify excessive billing by a single physician
  - Report entries against authorization records for new or terminated employees.
  - Identifying Multiple payroll deposited to same account number of the bank

There are various analytical techniques in existence for detecting fraud scheme in health care domain such as Calculation of statistical parameters, Stratification of numbers, Digital analysis using Benford's Law, Duplicate testing.

V. CHALLENGES IN HEALTH CARE DOMAIN

Though there are several advantages of using data analytics in health care domain, still challenges exist in this domain. The challenges in analytics includes data search, data capture, data storage, data sharing and data analysis. [5].

- ✓ *LACK OF ANALYTICS SKILL:* Healthcare providers are finding difficult to find the right specialists in analytics to help them get the best solutions from their databases. There is a shortage of data scientists having healthcare background to apply big data analytics to evaluate healthcare operations.
- ✓ *DATA SHARING:* Maintaining and centralizing Patient data records is effective within medical campus or hospital. Sharing of patient health records to outside environment is tedious and risky task. EHR data stores may be standardized which is adopted within the health care provider environment where as external component of health care system may use different systems and protocols. Sharing entire medical data records and managing integration of various medical data is still a challenging task.
- ✓ *LACK OF INTEGRATION BETWEEN CLINICAL AND ADMINISTRATION SYSTEMS:* Even internally, there often is an integration gap between patient care and administration. Medical records maintained by physicians and on the hospital floor must be reflected in accurate insurance claims and patient billing. The data management system has to be configured to ensure that treatment codes match and care given is accurately tracked for both administrative purposes and analytics.

Some of the challenges faced by the health care with big data analytics and how to overcome these challenges are listed in the Table 2.

Challenges	Solutions
Speed of browsing and processing the data	Use parallel processing or suitable Use Suitable hardware for fast scan of large amount of data
Data Security	Adherence to the standard guidelines and procedures Change of password frequently which only the user knows Upgrading Latest related security applications
Cyber-attack on medical devices	Uses Medical devices on secured Wi-Fi The devices are used for their intended purposes [6].
Large Volume of data from varied sources	Reliability of the data should be checked before arriving to a conclusion

Table 2: Big Data Analytics challenges and solutions in health care domain

VI. TRENDS IN BIG DATA ANALYTICS

- ✓ *VALUE-BASED, PATIENT-CENTRIC CARE:* The aim is to provide quality healthcare with coordination, support for payment structures and reduce healthcare costs.
- ✓ *HEALTHCARE INTERNET OF THINGS (IOT):* In order to offer more facilities at lower costs and improve patient care many of the IoT devices can be used to monitor the patient's behavior. These behaviors must be monitored by the physician at regular intervals of time. With the help of smarter IoT devices the patient monitoring process can be refined and properly medicated, thus improving the patient care by real time monitoring of patients.

Spending on healthcare IoT could top $120 billion in just four years, by some estimates. Most of the data created by the healthcare IoT is of the unstructured variety, creating a major role for Hadoop and advanced big data analytics working within the Hadoop framework.

- ✓ *Predictive Analytics for Outcomes Improvement– Predictive modeling and machine learning on large*

database samples, with more patient data, can find the patterns that couldn't be previously uncovered. This helps the doctors in making proper decisions based on the big-data to improve the treatment of patients. The machine learning algorithms help in identifying the various diseases at early stages, providing preventive measures and diagnosis to reduce the expensive complications.

Recently many companies are focusing on health care data analytics and contributing to the society providing efficient solutions and services to health care providers and insurers in order to have real-time monitoring of patients, improve health care, provides rich data for analysis, Intuitive care tools, and patient- centric decision support, reducing costs and increasing revenues [9 -10]. The Table 3 lists the companies with their contribution towards health care data analytics solutions.

Company Name	Description
Binary Fountain Software Next Health Technologies Inovalon	Provides SaaS Based Solution to extract real-time operating intelligence, control rich data, prescriptive analysis and behavioral economics to manage and optimize engagement of consumers
Health Catalyst SCIO Health Analytics	Provides process improvement and integrated data warehousing solutions to Improve Health Care, decrease costs, Increase revenues
Persivia, IBM, Media's Health, Optum Health, Truven Health Analytics	Provides Patient -Centric Clinical Decision Support Systems technology
Verisk Analytics RCG Global Services Mekesson	Real time monitoring and provides services to Health care providers and Insurers
Transcend Insight Green Way Health	Provides Intuitive Care Tools and Real-time Health Care Analytics
Vitreo Health	Provides advanced population health analytics solution

Table 3: List of Companies Providing Health Care Data Analytics Solutions

VII. CONCLUSION

The healthcare domain needs to gain the value added services of emerging technologies such as Data Science, Big Data, Analytics, Mobile computing and Cloud computing, along with the use of Health Information Technology (HIT) to help solve the ever growing operating cost problems employing healthcare analytics with efficient organization, streamlining and analysis of big data, will ensure prompt, precise and real-time diagnosis, decrease in preventable mistakes, and suitable treatment, thus proving to be more advantageous to the overall stake holders of healthcare system.

REFERENCES

[1] K. Rajalakshmi, Dr. S. S. Dhenakaran, "Analysis of Datamining Prediction Techniques in Healthcare Management System", International Journal of Advanced Research in Computer Science and Software Engineering Volume 5, Issue 4, April 2015

[2] Shankar Krishnan, "Application of Analytics to Big Data in Healthcare", 32nd Southern Biomedical Engineering Conference, pp 156-157, 2016

[3] Daniel M. Trifiletti, Timothy N. Showalter, "Big Data and Comparative Effectiveness Research in Radiation Oncology: Synergy and Accelerated Discovery", PMC Journal, Published online 2015 Dec 8.

[4] Coderre, David G., "Fraud Detection Using Data Analytics in the Healthcare Industry "Fraud Analysis Techniques Using ACL, John Wiley & Sons, 2009.

[5] Nina S Godbole , John lamb, "Using Data science and Big Data Analytics to make health care green", 12th International Conference & Expo on Emerging Technologies for a Smarter World (CEWIT) 2015, IEEE

[6] P. Mathew and D. Pillai, 'Big Data Solutions in Healthcare: Problems and Perspectives', IEEE, 2015.

[7] Manish Kumar Pandey and Karthikeyan Subbiah ,"A Novel Storage Architecture for facilitating Efficient Analytics of Health Informatics Big Data in Cloud , IEEE International Conference on Computer and Information Technology, 2016. PP 578 -585

[8] http://www.healthcareguy.com/2011/07/11/why-medical-device-data-is-the-best-way-to-fill- meaningful-use-ehrs-and-conduct-comparative-effectiveness-research-cer/

[9] http://www.healthcaretechoutlook.com/vendors/most-promising-healthcare-analytics-solution-providers-2016.html

[10] http://www.healthcaretechoutlook.com/vendors/most-promising-healthcare-analytics-solution-providers-2015.html