

Need, Application and Characteristics of Big Data Analytics in Healthcare-A Survey

M. Salomi^{1*} and S. Appavu Alias Balamurugan²

¹Department of CSE, Velammal College of Engineering and Technology, Madurai – 625009, India; msl@vcet.ac.in

²Department of IT, KLN College of Information and Technology, Madurai – 630612, India; app_s@yahoo.com

Abstract

Background/Objectives: To describe about the importance of big data analytics in healthcare, discuss the advantages and methodologies, describe examples provided in the literature, briefly discuss the challenges, and offer conclusion. **Methods/Statistical analysis:** A survey on methods that supports healthcare value by reducing cost at the same or better quality using big data techniques. **Findings:** Right procedure to guide the beginners of medicine and enhance Research and Development productivity in discovery, development, and safety in healthcare. It delivers the needed outcomes for each patient using evidence-based care, while ensuring safety. **Improvements/Applications:** Various approaches discussed can be applied in different fields of healthcare analytics in order to obtain sustainable answer for morbidity cases and also provide a better decision support system saving lives at lower costs.

Keywords: Big Data, Evidence Based Care, Genomic Analytics, Pre-Adjudication and Remote Patient Monitoring

1. Introduction

The large amounts of data, driven by record keeping, compliance and regulatory requirements and patient care has been historically generated in healthcare industry. The current healthcare system produces large volumes of distinct, structured and unstructured data. Thus, there is a prerequisite for a tool to collect and explore those large volumes of data. Big data analytics has contributed a lot for the evolution of healthcare research and become an efficient tool to manipulate large amount of clinical data in. Big data in healthcare is overwhelming not only because of its volume but also because of the diversity of data types and the speed at which it must be managed¹. The healthcare organizations will start realizing large benefits only by digitizing, combining and resourcefully using big data². Latent benefits include detection of healthcare fraud quickly and efficiently; detecting diseases at earlier stages more easily and effectively. Big data maps the diseases with symptoms and understands the patterns and learns the differences within the data, it has the talent to preserve lives with lower cost³. Big data analytics in healthcare is also suggested by various articles

like an *Evidence-based medicine*⁴ predicts patients at risk for diseases providing more efficient care and matching the treatments with outcomes by combining a variety of EMR data, operational and clinical data and analyzing them. *Genomic analytics*⁵ executes gene sequencing efficiently and cost effectively that provides regular medical care decision process. It makes genomic analysis a part of the regular medical care. *Pre-adjudication fraud analysis*⁶ hastily analyzes huge volumes of claim requests that reduces illegal hospital data. *Device/remote monitoring*⁷ is a revolutionary tele-health solution that captures and analyzes real-time large volumes of data and enables the transferring of information from the patient's medical device located outside the hospital to the clinical staff such as doctors and nurses in order to provide safety monitoring and adverse event prediction. Diverse stack holders of healthcare systems addresses the key pain points.

*Patient profile analytics*⁸ identifies patients who are similar to an index patient for decision support and Comparative Effectiveness Research (CER) analysis. This provides segmentation and predictive modeling that identifies the individuals, at high risk of getting affected by a specific disease, who would benefit from proactive

*Author for correspondence

care or lifestyle changes. This paper deals with the study of big data analytics in healthcare system. Firstly, we cover the various need, application and properties of big data analytics in healthcare. Secondly, four big data modeling techniques are described. Thirdly, the challenges are discovered. Fourthly, the examples of analytics of big data in healthcare reported in literature are provided. Lastly, conclusions are offered.

2. Need, Application and Features of Big Data Analytics in Healthcare

Healthcare organizations utterly depends on big data technology to capture all patient care data such as surgeon notes, lab reports, case history, X-Ray reports in order to get a more complete view of insight into care coordination and result based recompense models, fitness management, and patient commitment. Also collects the information such as list of doctors and nurses in a specified hospital, identifies the expiry date of the medicine and surgical instruments based on RFID data.

2.1 Need for Big Data Analytics

The quality of healthcare can be improved by considering the following: Providing patient centric services: Based on the available clinical data, easily detects diseases at earlier stages to provide faster relief to the patients. Uses evidence based medicine to minimize drug doses that help to provide efficient medicine and also avoid side effects based on patient's genetic composition¹. This reduces readmission rates and improves cost efficiency. Detecting spreading of diseases earlier: Based on the live analysis, the viral diseases are detected earlier before spreading. This can be identified by analyzing the patient's social logs suffering from a disease in a particular geo-location¹.

This helps the healthcare professionals to recommend the victims by taking required preventive actions. Monitoring the quality of hospitals: Monitoring the hospitals to ensure whether its setup is satisfying the norms setup by Indian medical council. This periodical check-up aids the government to take necessary actions against disqualifying hospitals. Improving the treatment methods: Based on the analysis of continuously monitoring the effect of medication, Physicians can change the dosages level in order to provide faster relief and provide proactive care to patients by monitoring patient vital signs

effectively. Analysis on the data generated by the patients, who already suffered from the same symptoms, helps physicians to provide effective medications to new patients.

2.2 Applications of Big Data

Big data is used in various fields like Device analytics capturing the vital signs from babies to detect advanced warning of the onset of complications. Clinical, financial and operational data for clinical decision support systems, transparency of medical data, aggregating and synthesizing patient, creates the reports they are required to create, analyzing their operational and financial data for efficiency gains are unified under Clinical Analytics. All patient related data (structured and unstructured) to get a 360-degree view of patient to measure and predict outcomes, manage patient population are unified under Outcome Analytics. Provider scoring and outcomes-based incentive calculation are provided to the prayer. Fraud Waste and Abuse Analytics Analyzing claims and benefits of OEF/OIF veteran's benefits and education fraud are addressed. Potential to do the above mentioned real-time application using Streams and big Insights. Drug discovery analytics integration of clinical, healthcare, patents, medical journals, compound information, public research data, and safety data to enable contextual, integrated access to correlated information around disease, target, and compound to provide key insights into decision making for target selection, compound selection, safety vs. efficacy issue discovery, lead optimization, clinical issue discovery, and so forth. Genomics Analytics combines patient genomic data along with the clinical data. Genomic data is becoming diagnostic to the complete patient record rather than an isolated data set which is self-sufficient.

2.3 Smarter Analysis using Big Data

Healthcare model that are driven by "Health outcomes" are called Evidence based Medicine. Inspection of structured and unstructured data Health Outcomes are required for analyzing care experience. Productivity of workforce and improvements in efficiency are highly demanded by Providers and shortage of staffs.

Looking at all the patient's data for providing optimal patient centered care which is implied by knowing patients. By knowing the lifestyle of the patients and their habits it is helpful to drive outcomes that are optimal and also becomes a part of providing comprehensive care and also

pre and post visits. In order to identify the chronic and reemerging diseases full scale processing of structured and unstructured information by the Disease Management that conducts management of diseases and inspection

3. Survey on Big Data Modeling Techniques

3.1 Evidence based Modeling

The best research evidence into the decision making process to take care of the patient, clinical expertise, values of the patient and the integration of all this is EBP. Clinical skills, education and the cumulated experience of the clinician is called clinical expertise. Every patient comes out with his or her preferences in person, their individual concerns, their expectations and values. A research evidence that is considered the best is generally the one found relevant in research that are clinical, and conducted by sound methodology that's represented in Figure 1.

Steps in EBP Process (Figure 2)

The feel that evidence based medicine at many condemnations core, it has been detached from the real-world practice of medicine. This view may have valid grounds. People suffering from single illness, with an extreme population of aging, and multiple morbidities are included in the trial. Evidence –based medicine promotes the following defined pathways. However, patients might not fit well with these pathways, and might not want to follow them.

Evidence-based medicine continues to produce a tidal wave of information. A quick and reachable answer

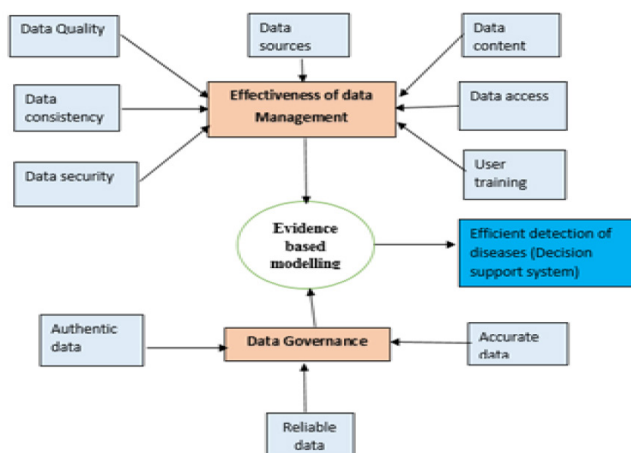


Figure 1. Evidence based modeling.

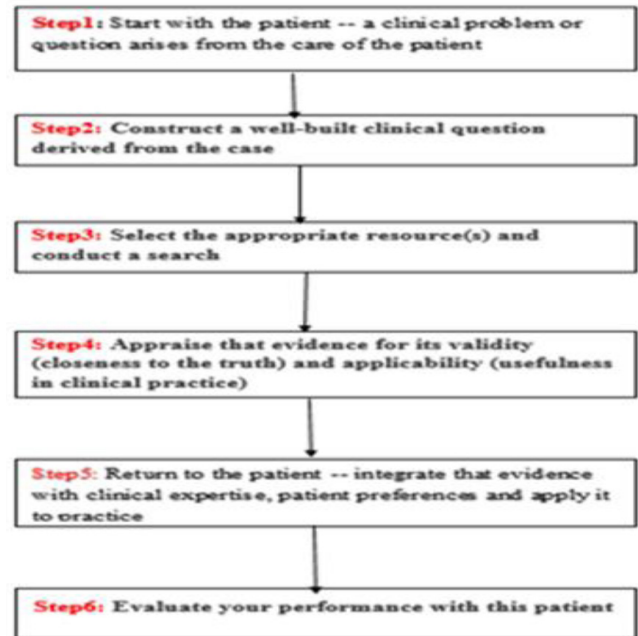


Figure 2. Steps in EBP process.

for an individual question are needed by a practicing clinician. In addition, trials that find statistically important outcomes can have very less clinical significance or applicability. It will have to address issues such as multiple morbidity; it will have to provide accessible answers; and it will have to deal in clinically important and patient-relevant outcomes as well as statistically significant ones.

3.2 Genomic Analytics

The role of genomics determines the impact of a particular drug on the patient. Genomics transition helps big data to transform into a day-to-day clinical application and explore new ways to adopt personalized medicine in a clinical setting. Big data capabilities and innovation techniques provides revolution in genomics that helps to reduce the cost of genome sequencing in the past two decades. Genome data analysis results in major developments in the personalized medicine field by developing treatments based on tumor patient's profile. The first step is to collect tumor samples from the cancer patients and to implement genome sequencing to understand relevant genetic abnormality in patients. Then, the eligible patients are allowed to join the clinical trial treatment and receive the drug to target the abnormality. Patient's diagnosis and treatment helps to make better decisions on genome sequencing derive.

3.2.1 Big Data in Enhancing the Decision

Sharing the genomic data on a common platform is the most important applications of big data analytics. In order to provide better result on treatment, variety of researches and clinical purposes are done with the shared data of each patient

Hence, for the privacy concern publishing and sharing the patient's data is restricted. Chances are available to overcome the limitations by confining data access to physicians and researches and also removing personally identifiable individual's information. For instance, the physicians be able to identify a set of patients suffered from similar symptoms of diseases and then analyze which treatment succeeded in that particular set of patients. After overcoming the technical challenges of big data processes, focus on the approach of transitioning research oriented of genomics into day-to-day clinical practice. Now-a-days, genomic testing is mainly implemented in big research and hospitals which have essential funds and resources that result in smaller population testing. Thus, there is a big need to reduce the investment and cost for testing which helps to maximize the patient pool and also to improve the adoption level of genomic technology in a clinical setting.

3.3 Pre-Adjudication

Inaccurate or fraudulent billing schemes are considered to be the most common process. Currently the existing data mining software is used to define behavior models that help to detect the common fraud and abuse schemes. Hence, there is increased risk that involves miscoded claim services which doesn't meet coding policies. Risk management also includes the ability to completely examine huge volume of claims from healthcare organizations, providers or services provided and the need to explore the outliers.

As the rules continuously change, payers need an automation solution for repairing select conditions which would otherwise lead to pending claims and require costly manual intervention. Advanced analytic systems enable identifying, predicting and minimizing fraud. Also checks the accuracy and consistency of claims. A large number of claim requests can be analyzed rapidly to reduce fraud¹. Pegasystem's Claims Repair, a combination of payer-specific business rules and claims processing best-practice templates, provides a solution. Softheon's BCB solution provides all necessary tools to the payers to deal with claims pre-Adjudication process that provides greater accountability, higher levels of accuracy, seamless

integration with leading technologies, interdepartmental communications and replacement of existing repositories.

3.4 Remote Patient Monitoring

Rapid changes in technology and the increase in mobile penetration shift the healthcare organizations towards an accountable care. Maximum number of senior citizens with lack of medical experts, all over the world, are the major reason for the need of remote tele-health services. In order to cut down healthcare costs, attain operational efficiencies and high quality accountable patient care, there is a need for wider acceptance of remote patient monitoring services⁹. Remote Patient Monitoring (RPM) is an emerging telehealth solution that captures and analyzes real-time large volumes of data and facilitates the exchange and flow of information from the patient's medical device located out of the hospital to the clinical staff such as doctors and nurses in order to provide safety monitoring and adverse event prediction. The healthcare system also addresses the key difficulties of diverse stakeholders. The patients are not forced to stay in the hospital for observation and they are also benefitted from round the clock high quality monitoring using RPM⁶. RPM mainly monitors the patient's condition; avoids medical emergencies and hospital re-admissions. It also helps setting of geographically isolated processes to access specialized and preventive treatments. Benefits of using Remote Patient Monitoring involve high savings for patients, health management, cutting healthcare expenses by billions¹⁰.

4. Challenges and Benefits

Lack of technically equipped persons in medical organizations. Relatively expensive for approaching accountable healthcare. Increase in severe medical conditions, senior citizen's population, diseases patterns and so on. Therefore, there is a need for best decision support system to predict the diseases at earlier stages¹¹. Increased rate for hospital readmission. Stress and strain combined with medical organizations and appropriate care. Additional pay-outs are caused due to the Increase Average Length Of Stay (ALOS) in the hospitals by insurance companies¹².

5. Conclusion

Large volume of data is being available in the real world almost in all applications. By effectively using big data

technologies, the required information can be extracted from the available large volume of data. Especially in healthcare organizations, the extracted information can be utilized to provide best decision support system that can save people life. Thus, the whole healthcare systems can stand to realize benefits from utilizing big data technologies. The throughput and reproducibility of methods for analyzing diseases has improved at a rapid rate, many challenges remain. For example, collecting accurate clinical information on samples remains an important and difficult task. However, available samples have lack of suitable consent to permit the linkage of clinical information. Future progress requires the expansion and improvement in infrastructure to collect samples and associated with suitable consent. Big data alone will not solve any issues for the health care problems that exist for individual patients and communities. Proper implementation of automation, analytics, and action, can help properly leverage big data for new solutions to health care models.

6. References

1. Sullivan F. Reducing information technology complexities and costs for healthcare organizations. *Drowning in Big Data*; 2014 Jul.
2. Raghupathi W, Raghupathi V. Big data analytics in healthcare: promise and potential. *Health Information Science and Systems*. 2014 Feb; 2(3):1–10.
3. Raghupathi W, Kudyba S. Taylor Francis. *Data Mining in Health Care*. In *Healthcare Informatics. Improving Efficiency and Productivity*; 2010 Apr. p. 211–23.
4. Big data Platform [Internet]. [cited 2012 May 1]. Available from: <https://www-01.ibm.com/software/data/bigdata>.
5. Case studies showcasing IBM client stories [Internet]. [cited 2012 Apr 22]. Available from: <http://www-03.ibm.com/software/businesscasestudies>.
6. Ohlhorst FJ. Professional speaker and IT business consultant; Wiley Publishers; 2012 Nov.
7. Dwoskin E, Walker J. Can data from your fit bit transform medicine? *The Wall Street Journal*; 2014 Jun 23.
8. Sun J, Hu J, Luo D, Markatou M, Wang F, Edabollahi Combining. *American Medical Informatics Association Annual Symposium (AMIA)*. 2012. p. 901–10.
9. TechMahindra [Internet]. [Cited 2016 Jan 05]. Available from: http://www.techmahindra.com/industries/Enterprise/hls/healthcare_provider/remote_patient_monitoring.aspx.
10. Patient monitoring, big data, and the future of healthcare [Internet]. [Cited 2014 Aug 1]. Available from: <http://www.wired.com/insightspatient-monitoring-big-data-future-healthcare>.
11. Joshi SG, Jayendran S, Dalal R. Transforming telecom business: scaling the shift using predictive analytics. *Indian Journal of Science and Technology*. 2015 Feb; 8(S4):34–43. doi: 10.17485/ijst/2015/v8iS4/60362.
12. Burghard C. Big data and analytics key to accountable care success. *IDC. Health Insights*; 2012.