Application of Big Data Analytics: An Innovation in Health Care

Swarna Latha Garapati [1], Dr. Sridevi Garapati [2]

[1]*Associate Professor, Dept. of Computer science, Rajamahendri Institute of Engineering and Technology, Rajahmundry, Andhra Pradesh,*

[2]*Assistant Professor, Dept. of Community Medicine, Rangaraya Medical college, Kakinada, Andhra Pradesh, India.*

Abstract

Recent rapid increase in the generation of digital data and rapid development of computational science enable us to extract new insights from massive data sets, in various disciplines, including internet business and finance. The amount of data being digitally collected and stored is vast and expanding rapidly. As a result, the science of data management and analysis is also advancing to enable organizations to convert this vast resource into information and knowledge that helps them achieve their objectives. Computer scientists have invented the term *big data* to describe this evolving technology. One of the most promising areas where big data can be applied to make a change is healthcare. Average human lifespan is increasing along world population, which poses new challenges to today's treatment delivery methods. The application of Big Data methods and techniques is growing quickly in the domain of clinical medicine and healthcare administration. Healthcare analytics have the potential to reduce costs of treatment, predict outbreaks of epidemics, avoid preventable diseases and improve the quality of life in general. In this article, we would like to address the need of big data in medicine and healthcare.

Keywords: Health Care, Big Data Analytics, Innovation, Application, Medicine

INTRODUCTION:

Since the mid 1980s, the world has experienced an unprecedented explosion in the capacity to produce, store and communicate data, primarily in digital formats. Simultaneously, access to computing technologies in the form of the personal PC, smart phone and other handheld devices has mirrored this growth. With these

enhanced capabilities of data storage and rapid computation as well as real-time delivery of information via the internet, the average daily consumption of data by an individual has grown exponentially. According to Asigra, a Cloud Backup company since 1986, a staggering 90% of the data in the world today have been created only during the last two years[1]; and it is predicted the worldwide number of Internet Protocol (IP) addresses will quadruple by 2015. Moreover, it is forecasted three billion people will be online creating close to eight zettabytes of data two years from now.[1]

Big Data refers to the fact that data today is often too large and heterogeneous and changes too quickly to be stored, processed, and transformed into value by previous technologies. The technological trends drive Big Data: business processes are more and more executed electronically, consumers produce more and more data themselves - e.g. in social networks cloud-based processing of petabytes of data can be achieved at a fraction of the time and cost when adopting big data processing methods.[2] Big Data has silently crept into our daily routines and, with continued development of cheap data storage and availability of smart devices both regionally and in developing countries, the influence of Big Data will continue to grow. This influence has also carried over to healthcare. There's a huge need for big data in healthcare, due to rising costs in nations like the United States. As McKinsey report states, "After more than 20 years of steady increases, healthcare expenses now represent 17.6 percent of GDP —nearly $600 billion more than the expected benchmark for a nation of the United State's size and wealth."[3] In other words, healthcare costs are much higher than they should be, and they have been rising for the past few years. Clearly, we are in need of some smart, data-driven thinking in this area.

In healthcare, the term Big Data typically refers to large quantities of electronic health record, administrative claims, and clinical trial data as well as data collected from smart phone applications, wearable devices, social media, and personal genomics services; Predictive Analytics refers to innovative methods of analysis developed to overcome challenges associated with big data, including a variety of statistical techniques ranging from predictive modeling to machine learning to data mining. Predictive analytics using big data have been applied successfully in several areas of medication management, such as in the identification of complex patients or those at highest risk for medication noncompliance or adverse effects.

This paper gives an overview of Big Data, sources of big data in health care and its advantages, potential challenges and the projected impact on the future of medicine in general.

BIG DATA IN THE HEALTHCARE SECTOR –
THE 21st CENTURY REVOLUTION

WHAT IS BIG DATA?

The term "big data" appeared in 1997, and was first used by few specialists. In 2001, reference to the data explosion phenomenon was made by using the "3Vs"[4] i.e. increasing *volume* (amount of data), *velocity* (speed of data's generation, integration, sharing, and processing), and *variety* (heterogeneity of data types and sources). These attributes have been widely adopted to define "big data", while in some other definitions a fourth "V" has been also employed referring to *veracity* (to address the quality of data and consequently the quality of evidence that can be derived from these data). In 2012, Gartner updated definition as follows: *"Big data is high volume, high velocity, and/or high variety information assets that require new forms of processing to enable enhanced decision making, insight discovery and process optimization."*[5] It was also observed that the term "Big Data" could mean different things to different groups of people. A report delivered to the U.S. Congress in August 2012 defines big data as "large volumes of high velocity, complex, and variable data that require advanced techniques and technologies to enable the capture, storage, distribution, management and analysis of the information."[6] Big data are data whose scale, diversity, and complexity require new architecture, techniques, algorithms, and analytics to manage it and extract value and hidden knowledge from it.[7] As the size of data increases above a critical point, quantitative issues of data are transformed into qualitative issues in the capture, processing, storage, analysis, and visualization of data.

Big data has been successfully used in astronomy (eg, the Sloan Digital Sky Survey of telescopic information), retail sales (eg, Walmart's expansive number of transactions), search engines (eg, Google's customization of individual searches based on previous web data), and politics (eg, a campaign's focus of political advertisements on people most likely to support their candidate based on web searches).

BIG DATA IN HEALTHCARE SECTOR:

Big data in healthcare refers to electronic health data sets so large and complex that they are difficult (or impossible) to manage with traditional software and/or hardware; nor can they be easily managed with traditional or common data management tools and methods.[8] Big data is characterized as extremely large data sets that can be analysed computationally to find patterns, trends, and associations, visualization, querying, information privacy and predictive analytics on large wide spread collection of data. Data generation in health care systems has now reached exabyte levels (1 exabyte = 1 billion gigabytes).[9,10] 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 managed.[8] The increased use of telehealth will further test the storage capacity for patient data and the innovative use of Google Glass by physicians will also add to the social and behavioral aspects of Big Data.[11] Electronic Medical Records (EMRs) contain a plethora of data, such as patient demographics, clinical and genomic data, and are known for assisting with the flow of health care. To manage this data, data analytics may be used to make it useful for retrieval. Hence the concept

of "big data" can be applied. Today they are seen as a way for performing large-scale and low-cost health care analysis and decision-making. On the other hand, increasing gap between healthcare costs and outcomes is one of the most important issues, and many efforts to fill this gap are under way in many developed countries. The gap between healthcare costs and outcomes was analyzed to be the result of poor management of insights from research, poor usage of available evidence, and poor capture of care experience, all of which led to missed opportunities, wasted resources, and potential harm to patients.

Benefits of health-related Big Data have been demonstrated in three areas, namely to 1) prevent disease, 2) identify modifiable risk factors for disease, and 3) design interventions for health behavior change.[12] Rumsfeld et al[13] summarized eight areas of application of big data analytics to improve healthcare: 1) predictive modelling for risk and resource use; 2) population management; 3) drug and medical device safety surveillance; 4) disease and treatment heterogeneity; 5) precision medicine and clinical decision support; 6) quality of care and performance measurement; 7) public health; and 8) research applications. Therefore, a pressing need to improve healthcare quality and patient outcomes, increasing data availability, and increasing analytic capabilities are the drivers of the big data era of healthcare.[13]

Big Data streams in health can be broadly summarized into three categories.[14] Traditional medical data is primarily originated from the health system (e.g. EMRs, personal and family health history, medication history, lab reports, pathology results), where the objective of these analyses is to derive a better understanding of disease outcomes and their risk factors, reduce health system costs, and improve its efficiency. "Omics" data refer to large-scale datasets in the biological and molecular fields (e.g. genomics, microbiomics, proteomics, and metabolomics), where the aim of these analyses is to understand the mechanisms of diseases and accelerate the individualization of medical treatments (e.g. "precision medicine").[11,15] There are various sources of medical big data, such as administrative claim record, clinical registries, electronic health records, biometric data, patient-reported data, the internet, medical imaging, biomarker data, prospective cohort studies, and large clinical trials.[13,16] However, sources of big data in healthcare can be broadly classified as-Internal (e.g., electronic health records, clinical decision support systems, CPOE, etc.) and external (government sources, laboratories, pharmacies, insurance companies & HMOs, etc.), often in multiple formats (flat files, relational tables, ASCII/text, etc.) and residing at multiple locations (geographic as well as in different healthcare providers' sites) in numerous legacy and other applications (transaction processing applications, databases, etc.). Sources and data types include: [6]

1. Web and social media data: Clickstream and interaction data from Facebook, Twitter, LinkedIn, blogs, and the like. It can also include health plan websites, smartphone apps, etc.

2. Machine to machine data: readings from remote sensors, meters, and other vital sign devices.

3. Big transaction data: health care claims and other billing records increasingly

available in semi-structured and unstructured formats.

4. Biometric data: finger prints, genetics, handwriting, retinal scans, x-ray and other medical images, blood pressure, pulse and pulse-oximetry readings, and other similar types of data.

5. Human-generated data: unstructured and semi-structured data such as EMRs, physicians notes, email, and paper documents.

For the purpose of big data analytics, this data has to be pooled. In the second component the data is in a 'raw' state and needs to be processed or transformed, at which point several options are available. The data stays raw and services are used to call, retrieve and process the data. Another approach is data warehousing wherein data from various sources is aggregated and made ready for processing, although the data is not available in real-time. Via the steps of extract, transform, and load (ETL), data from diverse sources is cleansed and readied. Depending on whether the data is structured or unstructured, several data formats can be input to the big data analytics platform. The most significant platform for big data analytics is the open-source distributed data processing platform Hadoop (Apache platform. It belongs to the class "NoSQL" technologies—others include CouchDB and MongoDB—that evolved to aggregate data in unique ways. Hadoop has the potential to process extremely large amounts of data mainly by allocating partitioned data sets to numerous servers (nodes), each of which solves different parts of the larger problem and then integrates them for the final result.[17,18,19,20]

PROCESSING OF BIG DATA:

More widely used big data processing methods are:-

Distributed big data processing: Conventional data processing techniques do not scale to meet processing requirements of big data. MapReduce is a distributed data processing method often used to process big data. This method adopts a two-step approach where the problem is first split into many homogenous sub-problems ('map' step) and then outputs of sub-problems are combined to generate the overall output ('reduce' step).[21] A cloud-based implementation of MapReduce has been used to analyse electrophysiological data in epilepsy clinical research.[22] This processing method is also commonly used in complex biological data processing operations such as genome sequencing which required large computation capabilities.[23] Evidence exists that big data processing outperforms conventional techniques to support the increased size of datasets.[24] Current statistical methods may indeed have limitations when handling the scale of datasets associated with big data.

Predictive analytics: This method of analysis uses various statistical and data mining techniques to analyse historical and present data in order to predict future outcomes. Predictive analytics is already demonstrating its usefulness in applications that enable a smarter prediction of health care outcomes by combining clinical, insurance and public datasets. It supports "intelligent case management" which involves the development of programs that can have a higher impact on patient behavior.[25,26]

There are evidences for the use of these methods in secondary care, for purposes ranging from reducing readmissions to predicting clinical outcomes in patients.[27,28,29]

Crowdsourcing: This involves recruiting large numbers of people who collaboratively collect, filter, and analyze large amounts of data for a common purpose. Using the internet as the medium of participation, thousands of people participate in completing a small part of a problem (often offered in multiplicity for the purpose of validation). Gamified approaches (i.e. regular tasks built as computer games) have been developed for aiding the diagnosis or the labelling of biomedical images.[30] In HIV research, crowdsourcing has been used to identify important genes using 500 billion subsets related to HIV biology.[31]

ADVANTAGES OF MEDICAL BIG DATA:

The potential value of medical big data has been demonstrated in: 1) The delivery of personalized medicine; 2) The use of clinical decision support systems such as automated analysis of medical images and the mining of medical literature; 3) Tailoring diagnostic and treatment decisions and educational messages to support desired patient behaviors using mobile devices; 4) Big data-driven population health analyses revealing patterns that might have been missed if smaller batches of uniformly formatted data had been analyzed instead; and 5) Fraud detection and prevention.[32] And also application of big data to medical care includes: detecting diseases at earlier stages when they can be treated more easily and effectively; managing specific individual and population health and detecting health care fraud more quickly and efficiently and other benefits like demonstrated in the areas of disease prevention, identification of modifiable risk factors for the disease and design interventions for health behavior change.[12] Certain developments or outcomes may be predicted and/or estimated based on vast amounts of historical data, such as length of stay (LOS); patients who will choose elective surgery; patients who likely will not benefit from surgery; complications; patients at risk for medical complications; patients at risk for sepsis, MRSA, C. difficile, or other hospital-acquired illness; illness/disease progression; patients at risk for advancement in disease states; causal factors of illness/disease progression; and possible co-morbid conditions (EMC Consulting). According to,[33] areas in which enhanced data and analytics yield the greatest results include: pinpointing patients who are the greatest consumers of health resources or at the greatest risk for adverse outcomes; providing individuals with the information they need to make informed decisions and more effectively manage their own health as well as more easily adopt and track healthier behaviors; identifying treatments, programs and processes that do not deliver demonstrable benefits or cost too much; reducing readmissions by identifying environmental or lifestyle factors that increase risk or trigger adverse events[34] and adjusting treatment plans accordingly; Improving outcomes by examining vitals from at-home health monitors; managing population health by detecting vulnerabilities within patient populations during disease outbreaks or disasters; and bringing clinical, financial and operational data together to analyze resource utilization productively and in real time.[33]

The big data in healthcare allows for:

1. Strategic planning and for Informed Strategic Planning - University of Florida made use of Google Maps and free public health data to prepare heat maps targeted at multiple issues, such as population growth and chronic diseases.

2. Electronic Medical Records (EMRs) - The most widespread application of big data in healthcare is EMRs maintaining. They contain a plethora of data, such as patient demographics, clinical and genomic data, and are known for assisting with the flow of health care, today they are seen as a way for performing large-scale and low-cost health care analysis and decision-making. Every patient has his own digital record which includes demographics, medical history, allergies, laboratory test results etc. Records are shared via secure information systems and are available for healthcare providers from both public and private sector. Every record is comprised of one modifiable file, which means that doctors can implement changes over time with no paperwork and no danger of data replication.

3. Real-time alerting - In hospitals, Clinical Decision Support (CDS) software analyzes medical data on the spot, providing health practitioners with advice as they make prescriptive decisions.

4. Predictive Analytics in Healthcare - To help doctors make big data-informed decisions within seconds and improve patients' treatment, particularly useful in case of patients with complex medical histories, suffering from multiple conditions and to predict new tools for screening of diseases.

5. Telemedicine - The term refers to delivery of remote clinical services using technology. The increased use of telehealth will further test the storage capacity for patient data. It is used for primary consultations and initial diagnosis, remote patient monitoring, and medical education for health professionals. Some more specific uses include telesurgery – doctors can perform operations with the use of robots and high-speed real-time data delivery without physically being in the same location with a patient. Such use of big data in healthcare can be linked to the use of predictive analytics. Patients can avoid waiting lines and doctors don't waste time for unnecessary consultations and paperwork. Telemedicine also improves the availability of healthcare as patients' state can be monitored and consulted anywhere and anytime.

6. Big data is helping to prevent unnecessary emergency room visits.

However, the use of Big Data has now reached all areas of healthcare, biomedical research and population health. Health services researchers can combine administrative and clinical databases to develop predictive models to improve health policy.[7] The pharmaceutical industry manages huge repositories of clinical and molecular data for rational drug design and pharmacogenomics approaches.[35] In population health, disease registries and data from clinical records are being used for measuring the impact of health interventions.[36] Biomedical researchers have access

to new sources of genomic data (e.g., microbiome, epigenomics) and can explore new hypotheses to understand the molecular causes of diseases.[37] For the big data scientists, there is vast amount of array of data and opportunity. The field of organ transplantation benefits from large, comprehensive, transplant-specific national data sets available to researchers. By discovering associations and understanding patterns and trends within the data, big data analytics has the potential to improve care, save lives and lower costs. Thus, big data analytics applications in healthcare take advantage of the explosion in data to extract insights for making better informed decisions.[38,39,40]

CHALLENGES FOR MEDICAL BIG DATA:

Health care, biomedical research and population health are generating massive, complex, distributed, and often dynamic sets of data, and that the size and complexity of this data will pose both challenges and opportunities to health organizations. Medical big data can be affected by several sources of uncertainty, such as measurement errors, missing data, or errors in coding the information buried in textual reports. Therefore, the role of the domain knowledge may be dominant in both analyzing the data and interpreting the results.[41] Other distinctive features of medical big data in analytic aspects includes the different types of patient characteristics, which sometimes may require weighting; the time structure, which may be an additional dimension; and treatment information, time point of treatment decision and change (i.e., time-dependent confounding).[42]

One of the biggest hurdles standing in the way to use big data in healthcare is how medical data is spread across many sources governed by different states, hospitals, and administrative departments. Integration of these data sources would require developing a new infrastructure where all data providers collaborate with each other. The healthcare industry has been slow to embrace Big Data due to the cost of adding analytic functions to existing EHRs, privacy issues, poor-quality data, and a lack of willingness to share data;[43] and also there are many methodological issues, such as data inconsistency and instability, limitations of observational studies, validation, analytical issues, and legal issue. The main challenges for big data are availability, ease of use, scalability, ability to manipulate at different levels of granularity, privacy and security enablement, and quality assurance.[6,18,44] To succeed big data analytics in healthcare needs to be packaged, so it is menu-driven, no lag between data collection and processing, user-friendly and transparent. Real-time big data analytics is a key requirement in healthcare. The dynamic availability of numerous analytics algorithms, models and methods in a pull-down type of menu is also necessary for large-scale adoption. The important managerial issues of ownership, governance and standards have to be considered. Health care data is rarely standardized, often fragmented, or generated in legacy IT systems with incompatible formats.[6] These great challenges are need to be addressed.

ETHICAL ASPECTS:

As the adoption of big data increases, there are increased concerns about the ethical use of health data. Privacy and security needs to be considered as a primary consideration of any big data solution in health care, and the necessary legislation needs to be adopted to ensure that big data is not misused. Effort should be taken to keep individual identities from being identified during big data processing workflows. However, there is the possibility that this may result in data duplication during the integration of big data sources and affect the usefulness of big data. Most online services are free and user agreements generally state that the owner of the service can use the data collected from the application. It is necessary to regulate data available online (especially health-related data). Intermediate processors that enforce governance restrictions may be an effective method to handle privacy and ethical concerns of big data.[45]

CONCLUSIONS:

Recognizing, understanding, and using Big Data in terms of scientific research and healthcare are necessary at this time in order to arrive at best evidence in a world of ever increasing data. The confluence of Big Data interpretations will continue given the proliferation of data from scientific led endeavors, accelerating healthcare innovations, and the rise of Big Data in higher education as a result of embedding technologies and the proliferation of e-Learning in higher education. Saving time, money and energy using big data analytics in healthcare is necessary- we can reduce readmissions, adverse events and treatment optimization for diseases affecting multiple organ systems. Big Data is Saving Lives. At the same time, medicine and healthcare is lagging behind in the adoption of Big Data approaches. Issues such as guaranteeing privacy, safeguarding security, establishing standards and governance, and continually improving the tools and technologies will have greater attention as big data analytics become mainstreaming. In the future there will be rapid, widespread implementation and use of big data analytics across the healthcare organization and the healthcare industry and also it has to address several challenges. Finally application of big data technology was inevitable and that the healthcare sector would be one of the sectors expected to be benefited the most from big data technology.

REFERENCES:

[1] Rasid Z. Big data infographic: What is big data? [Internet] 2013. [Cited March 17, 2014]. Available from: http://www.asigra.com/blog/big-data-infographic-what-big-data.

[2] de Lusignan S, Cashman J, Poh N, Michalakidis G, Mason A, Desombre T, et al.Conducting requirements analyses for research using routinely collected health data: a model driven approach. Stud Health Technol Inform 2012;180:1105–7. [PubMed]

[3]   Henke N, McKinsey, ' Data analytics: Changing the practice of medicine'

        https://www.mckinsey.com/industries/healthcare-systems-and-services/our-insights/the-big-data-revolution-in-us-health-care

[4]   Laney D. 3D data management: controlling data volume, velocity and variety. Available at: http://blogs.gartner.com/doug-laney/files/2012/01/ad949-3D-Data-Management-Controlling-Data-Volume-Velocity-and-Variety.pdf. Retrieved: May 2014.

[5]   Laney, Douglas. "3D Data Management: Controlling Data Volume, Velocity and Variety". Gartner.

[6]   IHTT . Transforming Health Care through Big Data Strategies for leveraging big data in the health care industry. 2013.

[7]   Bellazzi R. Big data and biomedical informatics: a challenging opportunity. Yearb Med Inform. 2014;9:8–13. doi: 10.15265/IY-2014-0024. [PMC free article] [PubMed][Cross Ref]

[8]   Frost & Sullivan: Drowning in Big Data? Reducing Information Technology Complexities and Costs for Healthcare Organizations.http://www.emc.com/collateral/analyst-reports/frost-sullivan-reducing-information-technology-complexities-ar.pdf

[9]   Newman HB, Ellisman MH, Orcutt JA. Data-intensive e-science frontier research.Communications of the ACM 2003:46(11):68–77.

[10]   Müller H, Hanbury A, Al Shorbaji N. Health information search to deal with the exploding amount of health information produced. Methods Inf Med 2012. Dec 4;51(6):516–8. [PubMed]

[11]   Paton C. Will Google Glass become standard eyeware for Doctors? [Internet] 2014. [CitedMarch17,2014]. Available from: http://www.healthinformaticsforum.com/forum/topic/show?id=2068976%3ATopic%3A149324&xgs=1&xg_source=msg_share_topic

[12]   Barrett MA, Humblet O, Hiatt RA, Adler NE. Big data and disease prevention: From quantified self to quantified communities. Big Data 2013;1(3):168–75. [PubMed]

[13]   Rumsfeld JS, Joynt KE, Maddox TM. Big data analytics to improve cardiovascular care: promise and challenges. Nat Rev Cardiol. 2016;13:350–359. doi: 10.1038/nrcardio.2016.42. [PubMed] [Cross Ref]

[14]   Swan M. The quantified self: Fundamental disruption in big data science and biological discovery." Big Data 2013. Jun;1(2):85–99. [PubMed]

[15]   Smith C. How many people use the top social media? [Internet] 2011. [Cited March 17, 2014]. Available from: http://expandedramblings.com/index.php/resource-how-many-people-use-the-top-social-media/#.UxvcnuddWL.

[16]   Slobogean GP, Giannoudis PV, Frihagen F, Forte ML, Morshed S, Bhandari

M. Bigger data, bigger problems. J Orthop Trauma. 2015;29(Suppl 12):S43–S46. doi: 10.1097/BOT.000000000000463. [PubMed] [Cross Ref]

[17]   Borkar VR, Carey MJ, Chen L. Big data platforms: what's next? ACM Crossroads.2012;19(1):44–49. doi: 10.1145/2331042.2331057. [Cross Ref]

[18]   Ohlhorst F. Big Data Analytics: Turning Big Data into Big Money. USA: John Wiley & Sons; 2012.

[19]   Zikopoulos PC, DeRoos D, Parasuraman K, Deutsch T, Corrigan D, Giles J. Harness the Power of Big Data. McGraw-Hill: The IBM Big Data Platform; 2013.

[20]   Zikopoulos PC, Eaton C, DeRoos D, Deutsch T, Lapis G. Understanding Big Data – Analytics for Enterprise Class Hadoop and Streaming Data. McGraw-Hill: Aspen Institute; 2012.

[21]   Schadt EE, Linderman MD, Sorenson J, Lee L, Nolan GP. Computational solutions to large-scale data management and analysis. Nat Rev Genet 2010. Sep;11(9):647–57.[PMC free article] [PubMed]

[22]   Sahoo SS, Jayapandian C, Garg G, Kaffashi F, Chung S, Bozorgi A, et al. Heart beats in the cloud: distributed analysis of electrophysiological 'big data' using cloud computing for epilepsy clinical research. J Am Med Inform Assoc 2014. Mar-Apr;21(2):263–71.[PMC free article] [PubMed]

[23]   Mudunuri US, Khouja M, Repetski S, Venkataraman G, Che A, Luke BT, et al.Knowledge and Theme Discovery across Very Large Biological Data Sets Using Distributed Queries: A Prototype Combining Unstructured and Structured Data. 2013. Dec 2;8(12):e80503. [PMC free article] [PubMed]

[24]   Simpson SE, Madigan D, Zorych I, Schuemie MJ, Ryan PB, Suchard MA. Multiple self-controlled case series for large-scale longitudinal observational databases. Biometrics2013. Dec;69(4):893–902. [PubMed]

[25]   Fox B. Using big data for big impact. How predictive modeling can affect patient outcomes. Health Manag Technol 2012. Jan;33(1):32. [PubMed]

[26]   Fox B. Using big data for big impact. Leveraging data and analytics provides the foundation for rethinking how to impact patient behavior. Health Manag Technol 2011. Nov;32(11):16. [PubMed]

[27]   de Lissovoy G. Big data meets the electronic medical record: a commentary on "identifying patients at increased risk for unplanned readmission". Med Care 2013. Sep;51(9):759–60. [PubMed]

[28]   Choi M, Lee J, Ahn MJ, Kim Y. Nursing critical patient severity classification system predicts outcomes in patients admitted to surgical intensive care units: use of data from clinical data repository. Stud Health Technol Inform 2013;192:1063. [PubMed]

[29]   Celi LA, Mark RG, Stone DJ, Montgomery RA. "Big data" in the intensive

care unit. Closing the data loop. Am J Respir Crit Care Med 2013. Jun 1;187(11):1157–60.[PMC free article] [PubMed]

[30]   Mavandadi S, Dimitrov S, Feng S, Yu F, Yu R, Sikora U, et al. Crowd-sourced BioGames: managing the big data problem for next-generation lab-on-a-chip platforms. Lab Chip 2012. Oct 21;12(20):4102–6. [PMC free article] [PubMed]

[31]   Bushman FD, Barton S, Bailey A, Greig C, Malani N, Bandyopadhyay S, et al.Bringing it all together: big data and HIV research. AIDS 2013. Mar 13;27(5):835–8.[PMC free article] [PubMed]

[32]   Roski J, Bo-Linn GW, Andrews TA. Creating value in health care through big data: opportunities and policy implications. Health Aff (Millwood) 2014;33:1115–1122.   doi: 10.1377/hlthaff.2014.0147. [PubMed] [Cross Ref]

[33]   IBM . Data Driven Healthcare Organizations Use Big Data Analytics for Big Gains.2013.

[34]   IBM . Harvard Medical School. 2011.

[35]   Scruggs SB, Watson K, Su AI, Hermjakob H, Yates JR, 3rd, Lindsey ML, Ping P. Harnessing the heart of big data. Circ Res. 2015;116:1115–1119. doi: 10.1161/CIRCRESAHA.115.306013. [PMC   free   article] [PubMed] [Cross Ref]

[36]   Sinha A, Hripcsak G, Markatou M. Large datasets in biomedicine: a discussion of salient analytic issues. J Am Med Inform Assoc. 2009;16:759–767. doi: 10.1197/jamia.M2780.[PMC free article] [PubMed] [Cross Ref]

[37]   Dinov ID. Methodological challenges and analytic opportunities for modeling and interpreting Big Healthcare Data. Gigascience. 2016;5:12. doi: 10.1186/s13742-016-0117-6. [PMC free article] [PubMed] [Cross Ref]

[38]   Ikanow: Data Analytics for Healthcare: Creating Understanding from Big Data.http://info.ikanow.com/Portals/163225/docs/data-analytics-for-healthcare.pdf

[39]   jStart . "How Big Data Analytics Reduced Medicaid Re-admissions." A jStart Case Study. 2012.

[40]   Knowledgent . Big Data and Healthcare Payers. 2013.

[41]   Bellazzi R, Zupan B. Predictive data mining in clinical medicine: current issues and guidelines. Int J Med Inform. 2008;77:81–97. doi: 10.1016/j.ijmedinf.2006.11.006.[PubMed] [Cross Ref]

[42]   Binder H, Blettner M. Big data in medical science--a biostatistical view. Dtsch Arztebl Int. 2015;112:137–142. [PMC free article] [PubMed]

[43]   O'Connor F. Health-IT early adopters well poised for big-data advances in clinical medicine. [Internet] 2013. [Cited March 17, 2014]. Available

from:http://www.computerworld.com/s/article/9238063/Health_IT_early_ado pters_well_poised_for_big_data_advances_in_clinical_medicine.

[44] Bollier D. The Promise and Peril of Big Data. Washington, DC: The Aspen Institute; 2010.

[45] Liyanage H, Liaw ST, de Lusignan S. Accelerating the development of an information ecosystem in health care, by stimulating the growth of safe intermediate processing of health information (IPHI). Inform Prim Care 2012;20(2):81−6. [PubMed]