Big Data in Healthcare Management: A Review of Literature

Senthilkumar SA[1], Bharatendara K Rai[2], Amruta A Meshram[2], Angappa Gunasekaran[3], Chandrakumarmangalam S[4]

[1]Department of Management, Pondicherry University, Pondicherry, India

[2]Charlton College of Business, University of Massachusetts Dartmouth, North Dartmouth, USA

[3]School of Business and Public Administration, California State University, Bakersfield, USA

[4]Department of Management Studies, Anna University Regional Campus, Coimbatore, India

Email address:
senthilkumar_sa@yahoo.co.in (Senthilkumar SA)

Abstract: A systematic literature review of papers on big data in healthcare published between 2010 and 2015 was conducted. This paper reviews the definition, process, and use of big data in healthcare management. Unstructured data are growing very faster than semi-structured and structured data. 90 percentages of the big data are in a form of unstructured data, major steps of big data management in healthcare industry are data acquisition, storage of data, managing the data, analysis on data and data visualization. Recent researches targets on big data visualization tools. In this paper the authors analysed the effective tools used for visualization of big data and suggesting new visualization tools to manage the big data in healthcare industry. This article will be helpful to understand the processes and use of big data in healthcare management.

Keywords: Big Data, Data Acquisition, Data Storage, Data Analytics, Data Visualization, Healthcare Management

1. Introduction

Healthcare Industry is one of the world's biggest and widest developing industries. During, the recent years the healthcare management around the world is changing from disease-centered to a patient-centered model [1] and volume-based to a value-based healthcare delivery model [2]. Educating the superiority of health care and decreasing the cost is a principle behind the developing movement toward value based healthcare delivery model and patient-centered care. The volume and demand for big data in healthcare organizations are growing little by little [3]. To provide effective patient-centered care, it is essential to manage and analyze huge health data. The outdated data management implements are not sufficient enough to analyze big data as variety and volume of data sources have increased in the past two decades. There is a need for new and innovative big data tools and technologies that can meet and exceed the ability of managing healthcare data [4]. Research study predictions on the worldwide big data expenditure in the healthcare business to progress towards Compound Annual Growth Rate (CAGR) of 42% during this years 2014-2019 [5].

The big data are used to predict the diseases before they emerge based on the medical records. Many countries' public health systems are now providing electronic patient records with advanced medical imaging media [6]. The practice of big data takes the prospective to encounter the upcoming market needs and trends in healthcare establishments [7]. Big data provides a great opportunity for epidemiologists, physicians, and health policy experts to make data-driven judgments that will eventually develops the patient care [8]. The authors have used Google trends for analyzing the 'big data in healthcare' between 2010 and 2015. The resulting graph is shown in figure 1.
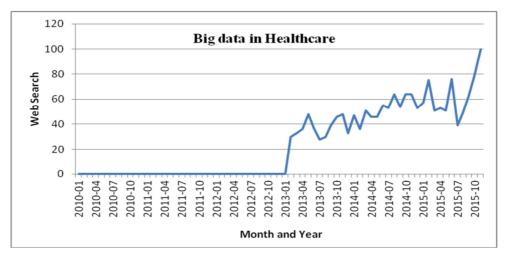
Figure 1. Google trend of big data in healthcare between 2010 and 2015.

Google trends is a unique, liberally available online portal of Google Inc. that permits consumers to cooperate with Internet hunt data, which may provide profound visions into peoples activities and health-connected occurrences. Google flu trends and Google dengue trends are used for nowcasting the spread of diseases like flu and dengue. Google trends have been used in several research publications [9].

It can be seen from Figure. 1 that the term 'big data in healthcare' really took off around early 2013. Increase in interest in this term can be related to a popular report by McKinsey & Company that came out in early 2013 [10]. The report highlights that healthcare expenses contributes about 17.6% of GDP and have a potential to reduce healthcare spending by $300 billion to $450 billion.

Currently, most big data studies in healthcare merely concentrate on technological understanding of big data [11]. Only a few topics had been debated the big data analytics in healthcare Information Technology [12]. The main concept of this is to review the processes and applications of big data in healthcare management.

This content is structured as follows: Section-2 introduction to the definitions and different dimensions of big data in healthcare. Section-3 discusses the search strategy and the steps in literature review of articles dealing with 'big data in healthcare'. In section-4, the process of big data management in healthcare is presented. Section-5 discusses the applications, benefits, and challenges of big data in healthcare. Since the privacy issues have been increasing recently, the laws related to protection and management of health and medical data are also discussed in this section. Finally the content finishes with conclusions and recommendations for upcoming investigation in the section-6.

2. Definition of Big Data

There are many definitions that have been provided by researchers for big data [13-14], however no sole definition of big data is generally known [15]. Baro et al [16] have done extensive research on the definition of bigdata and have proposed that a dataset could be qualified as "big dataset"

only if Log ($n*p$) is superior or equal to 7. Report submitted by U. S. Congress in August 2012 explains 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" [17]. In this content, we propose a new explanation of big data in healthcare.

The big data in healthcare involves collecting large collections of data from various healthcare foundations followed by storing, managing, analyzing, visualizing, and delivering information for effective decision making.

The big data in healthcare is associated with six characteristics viz., volume, variety, velocity, veracity, variability, and value. Various researchers have discussed first three V's (volume, variety, and velocity) that are widely used for characterizing big data in their papers [18-19]. The remaining three V's (veracity, variability, and value) are also discussed extensively by several researchers [20-21]. These six V's of big data are depicted visually in Figure 2.
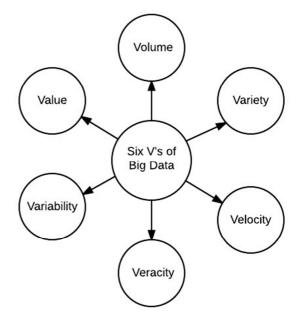

Figure 2. Six V's of big data.

Volume denotes to the large quantity of data produced by the organization. Today healthcare data are in terabytes (10^{12} bytes), petabytes (10^{15} bytes) or Exabyte's (10^{18} bytes)" [22]. In future the vast entirety of clinical data records will increase to zettabyte (10^{21} bytes) or yotta byte (10^{24} bytes). Such huge amount of data creates storage and massive analysis issue.

Variety arises from broadly contrasting bases of data or mash-ups of data resultant from autonomous sources in format [23]. The different format of data in healthcare can be classified as organized, semi-organized, or unorganized data [24]. Organized data includes laboratory data, clinical, sensor data and data from relational databases [25], semi- organized data includes data that is stored in Extensible Markup Language (XML) format [26], and unorganized data are free text data that usually does not have a precise design such as manual written note [27-28], data from X-ray images, radiological images and other medical imaging [29-30], Electronic Medical Record (EMR/HER) [31], graphics, patient discharge summaries, physiological measures (signals), healthcare data from social media and mobile phones [32-33]. 90 percent of big data are in the form of an unstructured data [34].

Velocity refers to the massive frequency during the current data is created, supplied and managed [35]. Velocity thus includes both equivalent the rapidity of data manufacture and the rapidity of data handling to meet demand. Accelerated increase of data is the third characteristic of big data [28, 29, 36]. The data that are generated can be either batch or real-time data [28]. The data's contents are frequently fluctuating through the concentration of corresponding data assemblies, the summary of previous data or inheritance gatherings, and the different forms of streamed data from multiple sources [25]. Example of velocity can be ageing of the population that is constantly increasing giving rise to number of patients, which increases growth rate of data by 55–60% every year [30].

Veracity refers to the correctness and accuracy of information [25]. Big data has low veracity, it can never be 100% accurate [37], and it is difficult to validate [38]. Since most of the data comes from unknown and unconfirmed sources, it is essential to set up a standard to ensure the feature of the data already it is involved.

Variability refers to data fluctuations throughout the handling and lifecycle. Developing range and variability also grows the attraction of data and the possibility in providing valuable information, unforeseen, and hidden [20].

Value is the method of extracting valuable information from huge sets of data and it is usually referred to as big data analytics [25]. Data value is useful for proper making decisions [39-21].

McKinsey & Company [10] believe that transforming of data deals with what is suitable or right healthcare environment and right for the for a patient They have considered following five strategic ways to significance, based on the concept that value is resulting from the balance of patient impact (outcomes) and. healthcare spend (cost).

Right living: Patients must be stimulated to be encouraged by taking an intimation part in their own health.

Right care: Patients must take the utmost timely, suitable treatment offered.

Right provider: Any professionals must have strong performance records who treats the patients and be skilled enough for succeeding the best results.

Right value: Suppliers and customers should frequently look for ways to expand value while maintaining or improving health-care quality.

Right innovation: Investors must concentrate on classifying new therapies and approaches to health-care delivery.

3. Methods

A complete literature analysis of article deals with 'big data in healthcare' related study were directed. A detailed search of publication of papers about big data in healthcare between 2010 and 2015 was used for review. The Table 1 includes the peer-reviewed journal articles from the large publishers' viz. Science Direct, PubMed, Springer, Taylor, & Francis, Inderscience and other reports on big data in healthcare.

Table 1. The article dealing with big data in healthcare from the peer-reviewed journals.

Name of the publishers	2010-11	2011-12	2012-13	2013-14	2014-15	Total 2010-15
Science Direct	1,323	1,615	1,772	2,262	2,645	9,617
PubMed	65	72	127	190	210	664
Springer	7	17	28	40	58	150
Taylor & Francis	0	0	0	9	42	51
Inderscience	0	0	3	4	2	9
Others	1	1		1	2	5
Grant total						10496

The detail search result yielded 10,496 papers. Our strategy for selecting the final review papers is explained in Figure 3. After reviewing these articles we discarded most of these papers.
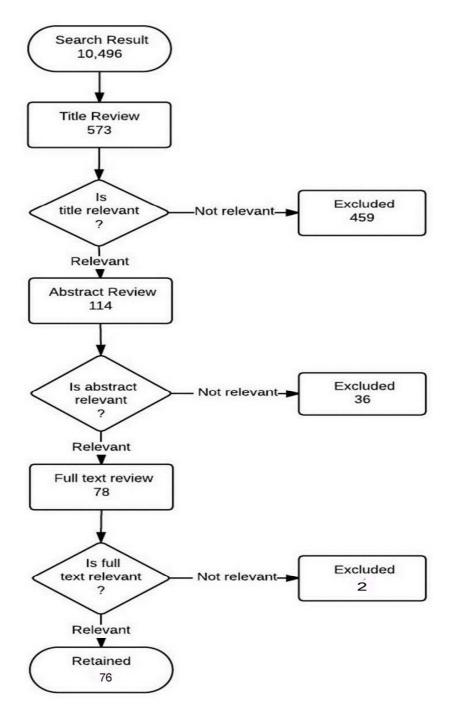
```
                    ┌──────────────────┐
                    │  Search Result   │
                    │     10,496       │
                    └──────────────────┘
                             │
                             ▼
                    ┌──────────────────┐
                    │   Title Review   │
                    │       573        │
                    └──────────────────┘
                             │
                             ▼
                         ╱Is     ╲
                        ╱ title   ╲────Not relevant──▶  ┌──────────────┐
                        ╲ relevant╱                     │  Excluded    │
                         ╲   ?   ╱                       │    459       │
                             │                           └──────────────┘
                          Relevant
                             ▼
                    ┌──────────────────┐
                    │ Abstract Review  │
                    │      114         │
                    └──────────────────┘
                             │
                             ▼
                         ╱Is       ╲
                        ╱ abstract  ╲───Not relevant──▶  ┌──────────────┐
                        ╲ relevant  ╱                    │  Excluded    │
                         ╲    ?    ╱                      │     36       │
                             │                            └──────────────┘
                          Relevant
                             ▼
                    ┌──────────────────┐
                    │ Full text review │
                    │       78         │
                    └──────────────────┘
                             │
                             ▼
                         ╱Is full   ╲
                        ╱ text       ╲──Not relevant──▶  ┌──────────────┐
                        ╲ relevant   ╱                   │  Excluded    │
                         ╲    ?     ╱                     │      2       │
                             │                            └──────────────┘
                          Relevant
                             ▼
                    ┌──────────────────┐
                    │    Retained      │
                    │       76         │
                    └──────────────────┘
```

Figure 3. *Review strategy.*

After removing irrelevant entries and duplications 573 papers were included for title review. After reading the title, 459 papers on big data that do not use the word healthcare in their title were excluded. Abstract reviewed by a human for eligibility, 36 papers were excluded since they were not directly related to healthcare and 78 papers relevant to healthcare full text criteria were reviewed. Finally the remaining 76 papers met all the inclusion criteria were retained. The Figure 3 shows the steps followed in search strategy.

4. Process of Big Data Analysis in Healthcare Industry

Big data analysis has the prospective to change the method of healthcare suppliers practice cultured equipment's to increase awareness from their clinical and other data repositories and make a declared conclusion. Big data healthcare analytics has five processes: Data Acquisition, Data Storage, Data Management, Data Analytics, and Data Visualization & Report. The Figure 4 presents the process of big data analysis in healthcare management.
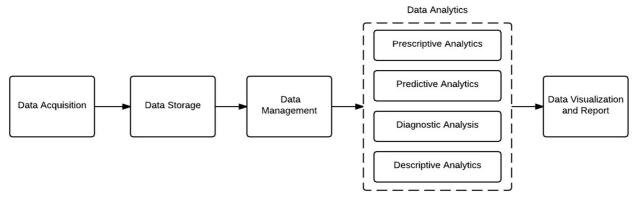
Figure 4. Big data in healthcare analysis process.

4.1. Data Acquisition

The big data in healthcare can be in a format of the structured, semi-structured or unstructured [40] and can be acquired from primary sources (e.g., CPOE, clinical decision support systems, electronic health records etc.) and secondary sources (laboratories, insurance companies government sources, pharmacies, & HMOs, etc.) [6] the following are the most important sources of big data in healthcare.

Electronic Health Records: During the earlier time span the digitization of healthcare histories have delivered a basis for hospitals on medical datasets [41]. Electronic Healthcare Records data from Physician Notes, Lab report, ECG, Scan, X-RAY, Health sensor devices details, Medical prescriptions etc. These set of data are the foundation for personalized medicine and large cohort studies [42] for hospitals.

Image Processing: Medical pictures are one of the cause of data for analysis. Computed tomography (CT), photo acoustic imaging, ultrasound, molecular imaging magnetic resonance imaging (MRI), mammography fluoroscopy, positron emission tomography-computed tomography (PET-CT), and X-ray are some of the examples of imaging procedures that are well recognized within clinical settings [43].

Social Media: Healthcare data can be collected from social media like Facebook, twitter, LinkedIn etc. Social Medical logs data are usually used for analyzing disease spreading/ transmission [44]. For example collecting information about a particular flu affected people from Twitter is faster than traditional method. Social networking cites like patients like me (www.patientslikeme.com) has more than 200,000 patients and is tracking more than 1,800 diseases [45].

Smart phones: Apps are the most important sources of data in the area of health care self-management. Nowadays, smart phones have health related apps like pedometers, fit bit that produces lot of data from number of stairs climbed, steps walked, and calories burned. Another app Mood panda is used to measure the individual mood and also anything from mental, emotional, and physical to social and environmental aspects of daily life. We also have apps for diabetes management I BG Star to monitor the blood glucose system. These apps create lot of data every day which contributes to healthcare research. Van Heerden et al. [46] used mobile phones to collect maternal health information from HIV-Positive Pregnant Woman in South Africa. Zhang et al. [47] showed that smart phones can be effectively utilized for domestic data gathering on infant feeding in rural china.

Web base data: The websites are also one of the most important source of healthcare data. The popular websites proving healthcare data are 23andMe and uBiome. 23andMe is a DNA study service providing evidence and implements for individuals to learn about and explore their DNA (https://www.23andme.com). uBiome is a Microbiome sequencing service that offers facts and tools for you to discover your micro biome (http://ubiome.com) [48].

4.2. Data Storage

The storage plays an vital part in big data. As a size of data in the healthcare industry is increasing we need an efficient and large storage platform. With such large cloud is the most promising technology. Clouds provides elasticity and proficiencies for get into data, creating awareness, accelerating the potential for scalable analytics solutions and driving value. Cloud computing is a powerful and promising technology to store enormous scale of data and perform enormous-scale and complex computing. It eradicates the need to sustain costly computing hardware, software and dedicated space.

Using cloud to examine big data in healthcare makes sense because:

Investments in big data investigation can be important and drive a need for efficient, cost-effective infrastructure.

Big data in healthcare organisations is a mix of internal and external foundations as they frequently keep patient's most delicate data in-house, enormous volumes of big data generated by public providers and third-party may be located outside.

Hashem et al. specified that cloud computing structure can assist as an actual platform to address the data storage essential to accomplish big data analysis [49].

Table 2 describes different platforms for cloud storage. It also describes their respective vendors and tools with purpose, advantages, disadvantages and application in healthcare.

***Table 2.** Cloud Storage Platforms in Healthcare.*

Types & Reference	Vendors	Components /platforms	Purpose	Advantages	Disadvantages	Application
Big data Storage [49]	Google Microsoft Amazon IBM	Cloud Services Azure S3 SmartCloud	The main purpose of is for saving large data and making easily accessible.	Enables the allocation of EHRs between certified physicians and hospitals in numerous Geographical areas, providing further appropriate contact to life-saving information and falling the need for replica testing.	1. Internet Bandwidth can be constraint if the speed is low. It may take longer time to upload files and retrieve them. 2. If no Internet, then no access to the storage 3. Data Security and Privacy is still a concern.	Cloud deployments for clinical applications with isolated or hybrid clouds provided these uses get maximum level of security, privacy, and availability. Nonclinical applications are a better fit for public arrangements but still must be wisely assessed.
Relational Databases [46]	Google Microsoft Amazon Cloudera IBM	Cloud SQL SQL Azure MySQl or Oracle MySQL, Oracle, PostgreSQL *dashDB*	They are fully succeeded SQL database service that makes it comfortable to set-up, maintain, bring about and administer relational MySQL databases in the cloud	Can access a familiar, highly available SQL database from Mapreduce applications, without having to worry about provisioning, management, and integration with other services.	There are connection, size, and App Engine specific limits in place.	Applications successively on Google App Engine or Google Compute Engine
NoSQL (Not only SQL) Database [7, 39, 49]	Google Microsoft Amazon Cloudera IBM	AppEngine Datastore Table storage DynamoDB Apache Accumulo *DB2*	speedy and elastic NoSQL database facilitates for all application need reliable, single-digit millisecond potential at any scale.	1. Flexible data models offered by NoSQL databases allow unstructured or semi-structured data to be stored easily 2. NoSQL databases are based on horizontal scalability which allows easy and automatic scaling	1. NoSQL database experts are difficult to get since the application itself is relatively new. 2. Though it is claimed that less administration is required, but they still require significant level of skill and effort to install and maintain.	Since the significance of EHR application for stability of care and complete health systems, using NoSQL databases have important possible to lead to superior EHR applications in terms of scaling, flexibility and greater obtainability.

4.3. Data Management

Data Management in healthcare includes organizing, cleaning, retrieval, data mining, and data governance. It also includes the method of validating whether there is some scrap data or any missing values. Such data needs to be removed [50]. It helps in risk assessment of patients, personalized discharge plan. Major data management tools are Apache Ambari and HCatalog. These tools are explained in detail in table 3. Data retrieval is a process of extracting file or valuable information from large healthcare databases. Big data analysis in healthcare frequently contains information recovery and data mining [51]. Wang et al. [52] mention that "information retrieval is the process of searching within large document collections, and in healthcare it mainly covers medical text retrieval and medical image retrieval. Data governance refers to overall management of security, integrity, usability and availability of the data employed in an enterprise. Maintaining confidentiality of individual patient records is very important in Healthcare management. This section provides information on key legislation to healthcare data accessibility and Government regulations intended to address health data privacy. The important data governance Act are HITECH, HIPPA, HDI, GINA and FOIA.

Health Information Technology for Economic and Clinical Health (HITECH) Act declared as portion of the American Recovery and Reinvestment Act of 2009, was engaged into law on February 17, 2009, to encourage the acceptance and significant use of health information knowledge. The implementation of HITECH Omnibus rule takes data management to a new level by mandating technological aspects of maintaining the patient record confidentiality. Subtitle D of the HITECH Act addresses the security and privacy concerns related with the electronic broadcast of health information, partly, through numerous requirements that reinforce the civil and illegal application of the HIPAA rules. HITECH involves all healthcare providers, to implement electronic healthcare system [53].

Health Insurance Portability and Accountability Act of 1996 (HIPAA) was considered to deliver secrecy of ethics to safeguard patients' health record and other medical information related to hospitals, doctors, health plans and other healthcare workers. HIPAA Title II also known as the administrative simplification provision which leads to the U. S. Department of Health and Human services to create national standards for handling electronic healthcare communications.

The HIPAA regulation is to declare that individuals' health information is appropriately secured however permitting the movement of health information required to deliver and promote high value of health care and to safeguard the populace's health and subsequent regulations direct covered entities (e.g., a healthcare institutions) in the safety of

individually recognizable secure the health information in research [54].

Department of Health and Human Services (HHS) under the Health Data Initiative (HDI) proposed a reliable privacy-preserved medical recommendation system. In this medical system, the patients can contribute their secured ratings of the physicians on different health conditions based on their satisfactions [55].

The Genetic Information Nondisclosure Act of 2008 (GINA) forbids health strategies and issuers from utilizing genetic data to make underwriting, premium-setting decisions, suitability or coverage [The Genetic Information Nondisclosure Act (GINA). Pub. L. No. 110–233, § 2000ff, 122 Stat. 881 (2008)]

The Privacy Act of 1974 & the United States Freedom of Information Act (FOIA) keeps recognizable facts about each personalities understood or gathered by the federal government [The Privacy Act. Pub. L. No. 93–579, § 552a, 88 Stat. 1896 (1974)].

4.4. Data Analytics

Data analytics is a process of transforming the raw data into information. Big data analytics in healthcare is classified into Descriptive, Diagnostic, Predictive, and Prescriptive Analytics [39]. The present research in academia and industry displays that retailers can attain up to 15 to 20% growth in ROI by setting big data into analytics [56].

Descriptive Analytics: It looks at past performance based on historical data. Descriptive Analytics is also known as unsupervised learning. It summarizes,

What happened in the healthcare management?

What is the impact of a parameter on the system?

Diagnostic Analysis: Using historical data predicts the root cause of problem and diagnose

Why did it happen?

Predictive Analytics: It analyzes both real-time and historical data, also known as supervised learning. It can only forecast what might happen in the future, because all predictive analytics are probabilistic in nature. It cannot predict the future. It anticipate

What will happen? What are the future trends?

What is the decision based on past history?

Prescriptive Analytics: This analytics automatically synthesizes big data and provide advice on number of different possible outcomes before the decisions are actually made. The decision maker can take this information and execute. Prescriptive analytics is advanced than the descriptive and predictive analytics. It prescribes

What should we do?

What is the best outcome and how can we make it happen?

Table 3 describes the tools which are being used to manage healthcare data. There are many analytical tools to analyze the data. We have classified the tools into 8 different layers. Each layer is further subdivided into components.

Table 3. Big data tools in healthcare.

Layers	Components	Description	Web Link	Review
Analytical Layer	Pentaho	Pentaho is use to design, develop and deploy Hadoop analytics solutions	http://www.pentaho.com	
	Apache Chukwa	Apache Chukwa is a data grouping method for handling huge circulated systems.	http://chukwa.apache.org/	[49]
Analysis Layer	Apache Pig	Apache Pig is a data-flow language in high-level, implementation of structure on behalf of similar calculation.	https://pig.apache.org/	
	Apache Spark	Apache Spark is use for processing large-scale data.	http://spark.apache.org/	[49]
	Cloudera Impala	Cloudera Impala is similar handling (MPP) SQL query engine for data storage in a computer cluster	http://www.cloudera.com/	
	Apache Tez	Apache Tez is a general data-flow programe structure, built on Hadoop	http://tez.apache.org/	
Abstraction Layer	Apache Oozie	Oozie is a workflow processing structure that lets users describes a sequence of works inscribed in different languages then wisely linked them to each other.	http://oozie.apache.org/	[7]
	Apache Cascading	Apache Cascading is use to generate and perform composite data processing workflows on a Hadoop cluster	http://www.cascading.org/	[24]
Processing Layer	Hadoop Mapreduce	A YARN-based system is processed parallel on huge data sets	https://hadoop.apache.org/docs/r1.2.1/mapred_tutorial.html	[7-34-20-39-57]
	Bidoop	Bidoop helps to analyze data sources and develop new analytical models	https://www.mapr.com/apps/bidoop-layer	
Integration layer (ETL)	Apache Flume	Apache Flume is use for mix together, combination, and moving huge volumes of log data	https://flume.apache.org/	[50]
	Apache Avro	Apache Avro is serialization framework	https://avro.apache.org/	
	Apache Sqoop	Apache Sqoop is a transitional instrument for data movement from non-Hadoop data into Hadoop data	http://sqoop.apache.org/	
Management Layer	Apache Ambari	Apache Ambari is a web-based tool for managing, monitoring, and provisioning Apache Hadoop clusters	https://ambari.apache.org/	
	Apache H Catalog	H Catalog is a centralized sharing service and metadata management	https://cwiki.apache.org/confluence/display/Hive/HCatalog+UsingHCat	
Coordinatio	Apache	A greater-presentation which coordinates service for distributed	http://zookeeper.apache.org/	

Layers	Components	Description	Web Link	Review
n Layer Storage Layer	Zookeeper	applications		
	Hadoop HDFS	Hadoop HDFS is a circulated file structure that delivers high-amount of contact on application data	https://hadoop.apache.org/docs/r1.2.1/hdfs_design.html	
	Apache Hive	Apache Hive is a data granary software abilities for managing and querying bulk datasets exist in circulated storage	https://hive.apache.org/	
	Apache HBASE	Apache HBASE is a dispersed database that enhances designed data storing for large tables	http://hbase.apache.org/	
	MongoDB	MongoDB is document-oriented database outlined	https://www.mongodb.org/	
	Apache Cassandra	Apache Cassandra is a scalable multi-master database with no opinions of disaster	http://cassandra.apache.org/	

4.5. Data Visualization

Data visualization is presenting the analytic results of healthcare data into pictorial or graphical format for understanding complex data and better decision making. It can be used to understand pattern and correlation among the data. The table 4 explains the tools that are currently being used for visualizing the big data in healthcare. The table describes description and features of the tool with its application in healthcare.

Table 4. Visualization tools used in healthcare.

Tool	Description	Features	Applications	References
R	R is a programming	- graphical display, control and manipulation of data	-R is been use by pharmaceutical company where it is used to plan clinical trials, and to forecast the finale day of the conclusions built on prearranged temporary studies of the data.	https://www.r-project.org/
	Language for advanced statistics and data visualization.	-Combined tools for instantaneous analysis	-Insurance concern practices R to build analytical models to set calculate risk profiles and premiums.	http://blog.revolution analytics.com/applications/
Cytoscape	Cytoscape is a common bioinformatics bundle for data integration and biological network visualization.	-Apps used for modelling advanced analysis -Visualize human-curated pathway datasets -Apps for problem domains	-Evaluation Model for Breast Cancer Vulnerability Gene --slim (InitGO) based on GO hierarchy structure and graph operations. -use to develop a semiautomatic technique to attain intact GO and self-governing	http://www.cytoscape.org/ Qijun Liu et al. Yan Jiang, Chao Xu
Graphviz	Graph visualization is a way of representing basic facts as figures of abstract, graphs and networks,	- flexible, -scalable -user-friendly software	-use for bioinformatics model	http://www.graphviz.org/
IBM Watson analysis	IBM Watson Analytics is a powerful analytics tool for cloud data	It can be use for predictive analytics, reporting, dashboards, visualization for large amount of data.		http://www-03.ibm.com/software/products/en/watson-analytics

The table contains 5 tools that can be used in future for visualizing healthcare data with their description, features and related link.

Table 5. Visualization tools can be used in healthcare.

Tools	Description	Features
Nodebox	It's a tool used for generative creative designs	-Data have been imported in various formats such as excel and in addition to that it is animation capable https://www.nodebox.net/
Flot	Flot is pure Javascript conspiracy collection for jQuery wit, attractive and collaborating, humble structures as focused.	-Helps in plotting categorical and textual data and usage of combinations for displaying elements in the same data series. -Produces interactive visualizations with toggling series http://www.flotcharts.org/
FF Chartwell	FF Chartwell evolution is a simple series of numerical into editable data concepts.	-For larger infographics the modules made are useful. -Use simple data series to produce charts and graphs https://www.fontfont.com/
Raphael	Raphael uses SVG and VML creating vector. A JavaScript library on the web for graphics, so that each representations formed is also a DOM object.	-Multi-chart skills -Create a range of graphs, charts and other data conceptions http://raphaeljs.com/
Cross filter	A JavaScript library that is skilled at controlling data sets by in excess of a million records. Discovering huge	-Discover huge multivariate datasets -Speedy incremental reducing and clarifying

Tools	Description	Features
Google Charts	multivariate data sets in a browser is made potential by Cross filter Google Charts offers fluctuating from basic distribute plots to categorised tree maps, a selection of data visualization formats.	-Increases presentation of live histograms http://square.github.io/crossfilter/ -Cross-browser compatibility -Cross-platform portability. https://developers.google.com/chart/
Nephi	Nephi allows users for difficult enquiry of links, and more for a better understanding of data relationships, and to both visualize and discover data social networks	- Real-time visualization -Deep data analysis to study relationships -In-built 3D version machine https://gephi.github.io/
Tableau Public	Tableau is a flexible-to-practice device for crafting, communicating, insert and data visualizations them on your website.	- Desktop presentation but accomplished graphics are kept on a public server - Drop and drag interface; -No programming skills required https://public.tableau.com/s/
Quadrigram	Quadrigram allows users to create entirely personalised visualizations using their own data and several mechanisms from a built-in library of entirety from charts and graphs to quadrification and loaded flow.	-Drafting the ideas and generating speed prototypes - Quick data processing by cloud-based computing system - Complete library of interactive visualizations -Build animations, dashboards and more http://www.quadrigram.com/
Perfuse	Perfuse is a data visualization tool which had been utilized by the IBM Visual Communication	-Data modeling, collaboration and visualization -Enhanced data arrangements for a diversity of pictorial designs http://prefuse.org/
Many Eyes	This tool offers a stand for generating a change of visualizations to demonstrate data point relationships,, create line and stack graphs, compare sets of values, examine text or view the many quantities of a total in a pie chart or tree map.	- Use existing data sets from other users or use your own -Displays data using Java and Flash -Upload datasets for public use -Full control to delete your data sets and visualizations--- -Choose from a multitude of ways to display data http://www-969.ibm.com/software/analytics/manyeyes/
iCharts	iCharts is a web-based presentation skilled of generating fascinating data visualizations for the purpose of web. Integrated graphs and charts into a website or application or allocate concluded visualizations through the iCharts Chart Channel. or social media	- Create interactive, employable charts -Add tags and descriptions for better discovery -Brand visualizations with your company logo - Allow social sharing -Qualified 3rd-party sites to re-embed visualizations to enlarged to spread http://icharts.net/
Visualize free	Based on the marketable visualization device InetSoft, Visualize Free is a free alternate for scrutinizing over numerous data sets and alternatives to recognise developments and operate data with a few simple clicks.	- Drag-and-drop constituents to figure the visualizations - Upload your data in CSV or Excel format - Share publicly or privately -Sandboxes for analysis and sales data http://visualizefree.com/

Table 6 describes all the tools that need programming knowledge to visualize healthcare data.

Table 6. Programming language supported visualization tools.

Name	Description	Features/ Related Links
D3. js	D3 is javascript library for creation of data visualizations.	- Numerous mechanisms and plugins are to develop abilities and the modules built in are reusable for the ease of coding. -Creates interactive SVG bar charts and bind arbitrary data to DOM - http://d3js.org/
Flare	An Action Script collection aimed at producing data visualizations that runs in Adobe Flash Player.	- Skilled of cooperating, difficult graphics -Maintenances visual encoding, data management, interaction and animation http://flare.prefuse.org/
SAS Visual Analytics	SAS Visual Analytics is the instrument to discover data sets of complete dimensions on visual basis for further widespread analytics.	- Drop and drag auto charting selects the best design for data. -It can be deployed on foundation or in private cloud or in public. - http://www.sas.com/en_us/software/business-intelligence/visual-analytics.html
Network X	Network X is skilled enough for producing graphs, digraphs and multigraphs based on datasets involved for multiple media formats. It based on the Python programming language	-Study the organization, dynamics and purposes of difficult networks http://networkx.github.io/

5. Applications of Big Data in Healthcare

Big data can be applied in almost all the areas of healthcare management. The potential application areas are fraud detection, epidemic spread prediction, Omics, clinical outcome, medical device design, insurance industry, personalized patient care and manufacturing, and

pharmaceutical development etc. [20]. Moreover the application of big data is widely adopted in personalized healthcare which offers an individual centric approach [58].

5.1. Applications of Big Data in 'Omics'

"Omics'' data refer to significant datasets in the organic and molecular fields (e.g., proteomics, metabolomics macrobiotics, genomics etc.). Application of big data on this study is to realize the strategies of diseases and increase the specification of medical treatments (e.g. "precision medicine") [59]. With the advance in metabolomics, proteomics, genomics, and other types of omics know-hows through the previous eras, a remarkable volume of data associated to molecular biology has been formed [60].

Genomics is the study of genes and their functions [61]. Application of big data in genomics will help to prevent or cure diseases and delivering personalized care to each patient [62]. This area is in still emerging period with presentations in particular concentrated regions, for example leukemia, diabetes, and cancer [63-64]. Pathway analysis is mostly used for high-quantity of genome-scale data [65], there are three generations of same structures used in pathway analysis [66]. The first generation tools are Clue Go, Onto-Express and GoMiner [67]. The most popular tool for second generation is GSEA [68], and the example for the third generation tool is Pathway-Express [69].

Proteomics is the study of proteome on their structures and functions. A proteome is the entire set of proteins in a cell. ExPASy (http://www.expasy.org/proteomics) lists dozens of databases on proteomics and over 100 tools. Big data application in proteomics, will have a major role in predicting and preventing human cancer [70]. Find Mod [71] and CSS-Palm [72] are frequently used for PTMs prediction.

Metabolomics is the systematic concept of chemical procedures including metabolites. The database BiGG used Genomic-based reconstruction of human metabolism for systems biology [73].

5.2. Insurance Industry / Payer

Healthcare Insurance companies/ payers are using big data in underwriting, fraud deduction, and claim management. Insurance providers are observing further than algorithmic fraud revealing practices that are claim-centric, to ones that are person–centric [74]. For example how many related claims were been submitted by the same personality or stated the identical treatment in different insurance companies.

5.3. Medical Device Design and Manufacturing

Big Data implement facilitates a wider set of device materials, delivery methods, and tissue interactions, anatomical configurations to be evaluated. Calculation techniques and Big Data can plays a significant role in medical system strategy and manufacturing [75].

5.4. Pharmaceuticals

Big data is used during all phases of pharmaceutical development, particularly for drug discovery [57]. Pfizer has recently initiated Precision Medicine Analytics Environment program that associates the dots among electronic medical record data, clinical trial, and genomic to identify chances to rapidly convey innovative medicines for particular patient populations.

5.5. Personalized Patient Care Healthcare

Big Data will make possible to bring best and modified patient care. In nearby future, fresh big data-derived influences will prompt suitable updates of diagnostic assistance, clinical guidelines and patient triage to permit more particular and modified treatment to advance medical result for patients (Yang et al., 2014)

6. Conclusions and Future Research Directions

This study reviewed the literatures on big data in healthcare. The works gap in the terms of method is recognized and connected with the empirical study.

The most challenge parts for big data in healthcare are data privacy, data leakage, data security, efficient handling of large volumes of medical imaging data, information confidentiality and security, wrong use of health data or failure to safegourd the healthcare information, and understanding unstructured clinical notes in the right context, extracting potentially useful information [76].

Some limitations have been identified after review, firstly there are limited publication that are available on big data in healthcare management.

This reading addressed the major experiments on data governance in healthcare. More detail study on data governance component can be future research topic.

The authors suggest few new data visualization tools to the healthcare analyst to make effective decision making. Big data has a great prospective to progress healthcare management and transform healthcare industry to next level. The review article will be benefiting the healthcare academicians, practitioners, researchers who are engaged in the areas of healthcare Management.

Conflict of Interests

The authors declared that there is no conflict of interests about the publication of this paper.

Acknowledgements

References

[1] J. W. Cortada, D. Gordon, B. Lenihan, The value of analytics in healthcare: From insights to outcomes, IBM Global Business Services, Executive Report, 2012.

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

[3] M. W. Stanton, Expanding patient-centered care to empower patients and assist providers, Research in Action. 5 (2002) 1-12.

[4] K. Feldman, D. Davis, N. V. Chawla, Scaling and contextualizing personalized healthcare: A case study of disease prediction algorithm integration, J. Biomed. Inform. (2015) 1–9.

[5] Global big data spending in the healthcare industry 2014-2019, Infiniti Research Limited. (Accessed on 2015, Sep 02). http://www.rnrmarketresearch.com/global-big-data-spending-in-healthcare-industry-2015-2019-market-report.html

[6] O. Y. Al-Jarrah, P. D. Yoo, S. Muhaidat, G. K. Karagiannidis, K. Taha, Efficient Machine Learning for Big Data: A Review, Big Data Res. 2 (2015) 87–93. doi:10.1016/j. bdr.2015.04.001.

[7] W. Raghupathi, V. Raghupathi, Big data analytics in healthcare: promise and potential, Heal. Inf. Sci. Syst. 2 (2014) 1–10. doi:10.1186/2047-2501-2-3.

[8] D. I. Sessler, Big Data and its contributions to peri-operative medicine, Anaesthesia. 69 (2014) 100–105.

[9] S. V. Nuti, B. Wayda, I. Ranasinghe, S. Wang, R. P. Dreyer, S. I. Chen, et al., The Use of Google Trends in Health Care Research: A Systematic Review, PLoS One. 9 (2014) e109583. doi:10.1371/journal. pone.0109583.

[10] P. Groves, B. Kayyali, D. Knott, S. Van Kuiken, The "big data" revolution in healthcare: Accelerating value and innovation, McKinsey & Company, 2013. http://www.pharmatalents.es/assets/files/Big_Data_Revolution.pdf.

[11] N. V. Chawla, D. A. Davis, Bringing big data to personalized healthcare: a patient-centered framework, J. General Internal Med. 28 (2013) 660-665. doi: 10.1007/s11606-013-2455-8.

[12] U. Srinivasan, B. Arunasalam, Leveraging big data analytics to reduce healthcare costs, IT Professional, 15 (2013) 21-28.

[13] A. D. Mauro, M. Greco, M. Grimaldi, What is big data? A consensual definition and a review of key research topics, 4th International Conference on Integrated Information, 2014.

[14] A. Gandomi, M. Haider, Beyond the hype: big data concepts, methods, and analytics, Int. J. Inf. Manage. 35 (2015) 137–144.

[15] J. D. Halamka, Early experiences with big data at an academic medical center, Health Aff. 33(2014) 1132–1138.

[16] E. Baro, S. Degoul, R. Beuscart, E. Chazard, Toward a Literature-Driven Definition of Big Data in Healthcare, BioMed Research International. 2015 (2015) 1–9. doi: 10.1155/2015/639021.

[17] IHTT: Transforming Health Care through Big Data Strategies for leveraging big data in the health care industry, 2013. http://ihealthtran.com/ wordpress/2013/03/iht%C2%B2-releases-big-data-research-report- download-today/.

[18] R. Leventhal, Trend: big data. Big data analytics: from volume to value, Healthcare Inform., Bus. Mag. Inf. Commun. Syst. 30 (2013) 12-14.

[19] Y. Wang, L. Kung, W. Y. C. Wang, C. G. Cegielski, Developing a Big Data-Enabled Transformation Model in Healthcare : A Practice Based View, in: Thirty Fifth Int. Conf. Inf. Syst., Auckland, 2014: pp. 1–12. doi:10.13140/2.1.2843.3601.

[20] L. Wang, C. A. Alexander, Big Data in Medical Applications and Health Care, Am. Med. J. 6 (2015) 1–8. doi:10.3844/amjsp.2015.1.8.

[21] X. Jin, B. W. Wah, X. Cheng, Y. Wang, Significance and Challenges of Big Data Research, Big Data Res. 2 (2015) 59–64. doi:10.1016/j. bdr.2015.01.006.

[22] M. L. Berger and V. Doban, Big data, advanced analytics and the future of comparative effectiveness research, Journal of Comparative Effectiveness Research. 3 (2014) 167–176.

[23] M. Sepulveda, Public health informatics and the public health workforce in an era of change, Am. J. Prev. Med. 47(2014) S386-S387.

[24] B. Feldman, E. M. Martin, T. Skotnes, Big Data in Healthcare - Hype and Hope, Dr. Bonnie 360 degree (Business Development for Digital Health), 2012. http://www.riss.kr/link?id=A99883549.

[25] G. Bello-Orgaz, J. J. Jung, D. Camacho, Social Big Data: Recent achievements and new challenges, Inf. Fusion. 000 (2015) 1–15. doi:10.1016/j. inffus.2015.08.005.

[26] A. O'Driscoll, J. Daugelaite, R. D. Sleator, Bigdata, Hadoop and cloud computing in genomics, Journal of Biomedical Informatics. 46 (2013) 774–781.

[27] N. M. S. Kumar, T. Eswari, P. Sampath, S. Lavanya, Predictive Methodology for Diabetic Data Analysis in Big Data, Procedia Comput. Sci. 50 (2015) 203–208. doi:10.1016/j. procs.2015.04.069.

[28] S. Bonney, HIM's role in managing big data: turning data collected by an EHR into information, Journal of American Health Information Management Association. 84-9 (2013) 62–64.

[29] J. D. van Horn, A. W. Toga, Human neuro imaging as a 'Big Data' science, Brain Imaging and Behaviour. 8 (2014) 323–331.

[30] M. Viceconti, P. Hunter, R. Hose, Big Data, Big Knowledge: Big Data for Personalized Healthcare, IEEE J. Biomed. Heal. Informatics. 19 (2015) 1209–1215. doi:10.1109/JBHI.2015.2406883.

[31] G. O. Matheson, M. Klugl, L. Engebretsen, Prevention and management of non communicable disease: the IOC consensus statement, Clinical Journal of Sport Medicine. 23 (2013) 419–429.

[32] K. D. Moore, K. Eyestone, D. C. Coddington, The big deal about big data, Healthcare Financial Management. 67 (2013) 60–68.

[33] T. H. Davenport, D. J. Patil, Data scientist: the sexiest job of the 21st century, Harvard Business Review. 90 (2012) 70–128.

[34] G. A. Ebenezer, S. Durga, Big Data Analytics in Healthcare: A Survey, ARPN J. Eng. Appl. Sci. 10 (2015) 3645–3650. doi:10.1155/2015/370194.

[35] K. Jee and G. H. Kim, Potentiality of big data in the medical sector: Focus on how to reshape the healthcare system, Healthc. Inform. Res. 19 (2013) 79–85. doi:10.4258/hir.2013.19.2.79.

[36] F. M. Afendi, N. Ono, Y. Nakamura et al., Data mining methods for omics and knowledge of crude medicinal plants toward big data biology, Computational and Structural Biotechnology Journal. 4 (2013) 1–14.

[37] J. C. Ward, Oncology reimbursement in the era of personalized medicine and big data, Journal of Oncology Practice.10 (2014) 83–86.

[38] O. S. Lupse, M. Crisan-Vida, L. Stoicu-Tivadar, E. Bernard, Supporting diagnosis and treatment in medical care based on big data processing, Studies in Health Technology and Informatics. 197 (2014) 65–69.

[39] S. Kaisler, F. Armour, J. A. Espinosa, W. Money, Big Data: Issues and Challenges Moving Forward, IEEE Comput. Soc. 46th Hawaii Int. Conf. Syst. Sci. (2013) 995–1004. doi:10.1109/HICSS.2013.645.

[40] R. Nambiar, R. Bhardwaj, A. Sethi, R. Vargheese, A look at challenges and opportunities of Big Data analytics in healthcare, Proc. - 2013 IEEE Int. Conf. Big Data, (2013) 17–22. doi:10.1109/BigData.2013.6691753.

[41] H. Chen, S. S. Fuller, C. Friedman, W. Hersh, Medical Informatics: Knowledge Management and Data Mining in Biomedicine, Springer Science & Business Media. 8 (2006).

[42] M. Stempniak, Beyond buzzwords: two state hospital associations collaborate around big data, Hosp. Health Netw. 88 (2014) 18.

[43] PatientsLikeMe. [Internet] 2015. [Cited, October 9, 2015]. Available from: https://www.patientslikeme.com

[44] A. Sadilek, H. Kautz, V. Silenzio, Modeling spread of disease from social interactions, Sixth AAAI International Conference on Weblogs and Social Media (ICWSM), 2012, http://www.cs.rochester.edu/~kautz/papers/Sadilek-Kautz-Silenzio_Modeling-Spread-of-Disease-from-Social-Interactions_ICWSM-12.pdf

[45] P. Wicks, M. Massagli, J. Frost, C. Brownstein, S. Okun, T. Vaughan et al., Sharing health data for better outcomes on PatientsLikeMe, J Med Internet Res. 12 (2010)e19. doi:10.2196/jmir.1549.

[46] A. van Heerden, S. Norris, S. Tollman, L. Richter, M. J. Rotheram-Borus, Collecting maternal health information from HIV positive pregnant women using mobile phone assisted face-to-face interviews in Southern Africa, J. Med. Internet Res. 15 (2013) e116.

[47] S. Zhang, Q. Wu, M. H. van Velthoven, L. Chen, J. Car, I. Rudan, Y. Zhang, Y. Li, R. W. Scherpbier, Smartphone versus pen-and-paper data collection of infant feeding practices in rural China, J. Med. Internet Res. 14 (2012) e119.

[48] M. Almalki, K. Gray, F. M. Sanchez, The use of self-quantification systems for personal health information: big data management activities and prospects, Heal. Inf. Sci. Syst. 3 (2015) 1–11. doi:10.1186/2047-2501-3-S1-S1.

[49] I. A. T. Hashem, I. Yaqoob, N. Badrul Anuar, S. Mokhtar, A. Gani, S. Ullah Khan, The rise of "Big Data" on cloud computing: Review and open research issues, Inf. Syst. 47 (2014) 98–115. doi:10.1016/j. is.2014.07.006.

[50] J. Archenaa, E. A. M. Anita, A Survey of Big Data Analytics in Healthcare and Government, Procedia Comput. Sci. 50 (2015) 408–413. doi:10.1016/j. procs.2015.04.021.

[51] L. Z. Andrzej Chluski, The application of big data in the management of healthcare organizations: A review of selected practical solutions, Bus. Informatics. 1 (2015) 9–18. doi:10.15611/ie.2015.1.01.

[52] L. Wang, R. Ranjan, J. Kołodziej, A. Zomaya, L. Alem, Software Tools and Techniques for Big Data Computing in Healthcare Clouds, Futur. Gener. Comput. Syst. 43 (2015) 38–39. doi:10.1016/j. future.2014.11.001.

[53] U. S Government, Department of Health and Human Services, Fedral Register, Rules and Regulations, 74(2009) 56123-56131, Available from: https://www.hhs.gov/sites/default/files/ocr/privacy/hipaa/administrative/enforcementrule/enfir.pdf

[54] Atchinson, Brian K., Fox, Daniel M. (May–June 1997), The Politics Of The Health Insurance Portability And Accountability Act, 16 (3): 146–15 doi: 10.1377/hlthaff.16.3.146.

[55] Martin Wiesner, Daniel Pfeifer, Health Recommender Systems: Concepts, Requirements, Technical basics and Challenges, International Journal of Environmental Research and Public Health, 11(2014) 2580-2607. doi:10.3390/ijerph110302580.

[56] J. Perrey, D. Spillecke, A. Umblijs, Smart analytics: How marketing drives short-term and long-term growth. McKinsey Quarterly (2013).

[57] T. Schultz, Turning healthcare challenges into big data opportunities: A use-case review across the pharmaceutical development lifecycle, Bull. Association Inform. Sci. Technol. 39 (2013) 34-40. doi: 10.1002/bult.2013.1720390508.

[58] Rohil Shah, Ria Echhpal, Sindhu Nair, Big Data in Healthcare Analytics, International journal on Recent and Innovation trends in Computing and Communication., 10(2015) 134-138.

[59] M. M. Hansen, T. Miron-Shatz, a Y. S. Lau, C. Paton, Big Data in Science and Healthcare: A Review of Recent Literature and Perspectives. Contribution of the IMIA Social Media Working Group, Yearb. Med. Inform. 9 (2014) 21–6. doi:10.15265/IY-2014-0004.

[60] S. Li, L. Kang, M Zhao X, A survey on evolutionary algorithm based hybrid intelligence in bioinformatics, BioMed Research International. 2014 (2014) 8. doi: 10.1155/2014/362738.362738.

[61] Genomics and World Health: Report of the Advisory Committee on Health research, Geneva, WHO (2002), http://www.who.int/genomics/geneticsVSgenomics/en/.

[62] A. Belle, R. Thiagarajan, S. M. R. Soroushmehr, F. Navidi, D. A. Beard, K. Najarian, Big Data Analytics in Healthcare, (n. d.). doi:10.1155/2015/370194.

[63] E. M. van Allen, N. Wagle, M. A. Levy, Clinical analysis and interpretation of cancer genome data, Journal of Clinical Oncology. 31 (2013)1825–1833.

[64] F. Andre, E. Mardis, M. Salm, J. C. Soria, L. L. Siu, C. Swanton, Prioritizing targets for precision cancer medicine, Annals of Oncology. 25 (2014) 2295–2303.

[65] D. W. Huang, B. T. Sherman, R. A. Lempicki, Bioinformatics enrichment tools: paths toward the comprehensive functional analysis of large gene lists, Nucleic Acids Research. 37 (2009) 1-13.

[66] Khatri, M. Sirota, A. J. Butte, Ten years of pathway analysis: current approaches and outstanding challenges, PLoS Computational Biology. 8 (2012) Article ID e1002375.

[67] S. Draghici, P. Khatri, R. P. Martins, G. C. Ostermeier, S. A. Krawetz, Global functional profiling of gene expression, Genomics. 81 (2003) 98–104.

[68] A. Subramanian, P. Tamayo, V. K. Mootha, Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles, Proceedings of the National Academy of Sciences of the United States of America. 102 (2005) 15545–15550.

[69] S. Draghici, P. Khatri, A. L. Tarca, A systems biology approach for pathway level analysis, Genome Research. 17 (2007) 1537–1545.

[70] H. K. Yalamanchili, Q. W. Xiao, J. Wang, A novel neural response algorithm for protein function prediction, BMC Syst. Biol. 6 (2012) S19.

[71] M. R. Wilkins, et al., High-throughput mass spectrometric discovery of protein post-translational modifications, J. Mol. Biol. 289 (1999) 645–657.

[72] J. Ren, et al., CSS-Palm 2.0: an updated software for palmitoylation sites pre- diction, Protein Eng. Des. Sel. 21 (2008) 639–644.

[73] N. T. Issa, S. W. Byers, S. Dakshanamurthy, Big data: the next frontier for innovation in therapeutics and healthcare, Expert Rev. Clin. Pharmacol. 7 (2014) 293–298. doi:10.1586/17512433.2014.905201.

[74] P. Bharal, A. Halfon, Making Sense of Big Data in Insurance, ACORD and MarkLogic, 2013.

[75] A. G. Erdman, D. F. Keefe, Grand challenge: Applying regulatory science and big data to improve medical device innovation, IEEE Ttrans. Biomed. Eng. 60 (2013) 700-706. doi: 10.1109/TBME.2013.2244600.

[76] R. Bellazzi, Big data and biomedical informatics: a challenging opportunity, Yearb. Med. Inform. 9 (2014) 8–13.