Global Journals LaTeX JournalKaleidoscope™
Artificial Intelligence formulated this projection for compatibility purposes from the original article published at Global Journals. However, this technology is currently in beta. *Therefore, kindly ignore odd layouts, missed formulae, text, tables, or figures.*

Potential of Big Data Analytics in Bio-Medical and Health Care Arena: An Exploratory Study

Mandava Geetha Bhargava[1], V. Sucharita[2] and P.Venkateswara Rao[3]

[1] KL University

Received: 12 December 2016 Accepted: 4 January 2017 Published: 15 January 2017

Abstract

With the leveraging emerging big Data in every industry, Big Data can amalgamate all data related to patient to get a complete view of patient to analyze and predict the outcomes. Using big data analytics as tools. It can enhance development in new drugs, health care financing process and clinical approaches and extends a lots of benefits such as better health care quality and efficiency, fraud detection and early disease detection by means of analytics of big data. This paper provides a general survey of current progress and advances in research arena of big data, bio-medical and health care and some major challenges of big data concept and characteristics, this concerns includes big data from bio-medical and health care arena, benefits of big data, its applications and opportunities, Methods and technology progress about big data in bio-medical and health care and challenges of big data in both bio-medical and healthcare are also discussed.

Index terms— big data; health care; genomics; big data analytics

1 I. Introduction

ig Data can be termed as massive data or complex that exceeds the processing capacity of traditional data processing applications and challenges are Acquire, Process, Manage, Generate, Capture, storage, sharing and visualize. Now a days Big Data is processed for analytics of various parameters in each and every field of work like Research, Education, I.T, Banking, Bio-Medical, Health Care, Construction, and Manufacturing etc. With help of some big data technologies it is been processed and characteristics of big data are 6V's i. wearable or implantable sensors like Bio-Metric, Blood Pressure and ECG etc. is gathered and analyzed in real time. The data in bio-medical and health care arena can be differentiate as follows: genomic data where it consists of DNA expression, genotyping and gene expression etc., clinical data where it consists of structured and unstructured data such as X-ray images, test reports, patient discharge reports and laboratory data etc., behavior and patient sentiment data consists of web data, social media data, streamed data such as telehealth and endoscopy etc., clinical reference and health publication data consists of practice guidelines, journals , medical reference material and health products etc., business, administrative and external data consists of financial data ,scheduling, billing and biometric data and other important data such as patient feedback and device data etc. as shown in figure.2. Big data can avail support over all aspects of biomedical and health care. Big data analytics has gained traction in analytics of fraud detection and prevention, clinical outcome, genomics, epidemic disease prediction, pharmaceutical development and personalized patient care, etc. there are potential applications in biomedical and health care as they are discussed below:

2 a) Fraud detection and prevention analytics

Detecting, predicting and reducing fraud can be executed by using advanced analytics systems for fraud detection and checking the consistency and accuracy of claims. Big data predictive modelling can be used by health care users and payers for fraud prevention.

3 b) Clinical outcome Analytics

Clinical analysis can be implemented through merging financial, operational and clinical data for efficient clinical assessments. Clinical data can be utilize to reduce the manage and predict the health risks and improve clinical outcomes with cost of care.

4 c) Genomics analytics

The data about genes and DNAs can be analyzed for predicting and reducing the rate of disease of patients and it is becoming critical to the complete patient record to merge the both genomics data and clinical data helps to cure perilous diseases such as cancer, etc.

5 d) Epidemic disease prediction

In public and population health, continuously analyzing and aggregating public health data helps identifying and managing potential disease out breaks by means of analytics of social media and web-based data the disease outbreak can be known based on social content ,query activity and consumer search.

6 e) Design and manufacturing of medical devices

Tools of big data allows a broader set of device materials, tissue interactions, delivery methods, and anatomical configurations to be analyzed. Big data and computational methods can play an important role in design and manufacturing of medical devices.

7 f) Pharmaceuticals development

By analytics of pharmaceuticals data, the pharmaceutical companies can increase their ability to continue bringing new life enhancing medicines to patients in a timely manner, on basis of management of big data which was generated during all phases of pharmaceutical development, the cost of pharmaceutical product will be cost effective.

8 III. Research Background

Kiyana zolfaghar et al [1] done research on solutions for predicting risk of readmission for congestive heart failure patients by means integrating data of national impatient dataset and patient dataset and developed a datamining predictive model by means of integrated data and concluded that effectiveness of quality, scalability, efficiency by means of big data infrastructure on the predictive model.

Sean D.Young et al [2] done research on approaches of utilising real time social media technologies for identification and remote monitoring of HIV outcomes through negative binomial regression of tweets and concluded that the feasibility of using real time social media data to detect HIV risk-related communications, geographically map the location of those conversations and link them to national HIV outcome data for additional analysis Priya Nambisan et al [3] done research on ruminating behavior of depression through social media, big data and public health informatics through tweets from micro blogging sites by means of screening the Year 2017 Volume XVII Issue II Version I vocabulary of tweets and shows sleep, pain and suicidal thoughts as they do offline and concluded that the characteristic can be used to detect and diagnose depression using the tweets in a much more effective and efficient way.

Zhendong Ji [4] done study on analysis of big data application in the medical industry and potential of its commercial value for the health care industry and concluded that by big data analysis in the medical industry provides future and promoting continuous development of medical industry through meta-analysis of gathered data.

Quan Zou et al [5] done a study on map reduce frame operation in bioinformatics through different applications and mechanisms of MapReduce and concluded that Hadoop framework has capable of handling bioinformatics data and traditional bioinformatics resources will be redesigned to support Hadoop MapReduce for high performance computing.

Liang y et al [6] studied on big data science and its applications in health care and medical research and concluded that big data offers new opportunities and promising with challenge in every field .the collaborative network, nurturing environments, team science approach with highly trained with computational skills, domain/disease expert and interdisciplinary are crucial.

Lidong Wang et al [7] done a study on big data in medical applications and health care by means of big data concept and characteristics, health care data and major issues of big data and concluded that big data can improve the research and development, translation of new therapies and has great potential to improve medicine, guide clinicians in delivering value based care.

9 IV. Methods and Technology Progress in Big Data

In health care /Bio-medical arena, massive amount of data about patient's medical histories, diagnosis and responses, symptomatology to treatments and therapies are gathered. Data mining techniques can implemented to derive knowledge from the gathered data in order to either examine reporting practices or to identify new

patterns in infection control data. Moreover, predictive models can be utilized as detection tools can be utilize as electronic patient record gathered for every individual person of the area.

Visual analytics presents a new area of research with big data by conceptualizing the output of complex processes. The appropriate visualization solutions to the big data examples such as real time interactive visualization and metrics dash boarding [9]. Unstructured data can be converted into form of tables is to put attributes of exchangeable image file (EXIF) tags or place analyzed data where it leads to easier at visualization process. Big data can be processed through cloud technologies where it provide us operationally, insights-clinically and in research [10]. The concept of STAAS (storage as a service), is a one of the services provided by cloud computing, which provides health care center with a massive amount of storage for processing on basis of demand at low cost. [11]. Beside general infrastructure of cloud (compute, storage, virtual machine management), the following services are required to handle big data [12]. Hadoop related frameworks and tools, specialist data analytics tools, Cluster services, massively parallel processing databases, databases /servers SQL, NoSQL and security infrastructure. Organizations used various methods to de-identification of the distance data from personal identities and preserve individual's privacy. Deidentification has been seen as an important security measure to be taken under the data security and accountability principle [13].

10 V. Proposed Approach and System

From studies of literature, the mechanisms and methodologies are basis upon Hadoop-MapReduce Framework in above literatures where it does only analysis and processing, having disadvantage of statistics analysis where it can't does and can't stores the data. Authors proposing a system which consists of RHadoop platform, it contains both R Language and Hadoop-MapReduce framework where it can process the different types of data with statistical data and streams the data after and before analysis through separate statistical package. It can visualize the processed data as output and figure 4 shows the architecture of Hadoop and data analysis tool. RHadoop is an open source project developed by revolution analytics, provides client side integration of Hadoop and R. It allows running a MapReduce jobs within R. Needs some packages for integration of R and For big data healthcare/biomedical systems, the combination of Hadoop-MapReduce framework and R language is uniquely capable of storing and analyzing wide range of healthcare data types including genomic data, financial data, electronic medical records and claims data etc. the combination of two frameworks offers availability, high scalability, statistical analysis and reliability than traditional data processing systems. In addition, intelligent functional components can be built such as surveillance, detection, notification, diagnosis and recognition etc. figure 3 shows a general framework of big data analytics. The statistical software which handles data to analyze statistically through by means of visualization of graph. The software's such as SAS, R and etc.

11 VI. Challenges of Big Data in

Biomedical and Health Care

Variety, volume and velocity characteristics of big data have brought challenges in retrieval, data storage ,visualization and search .veracity and variability of big data indicate data uncertainty and instability, which often makes big data analytics difficult and major challenges of big data in bio-medical and health care are as follows:

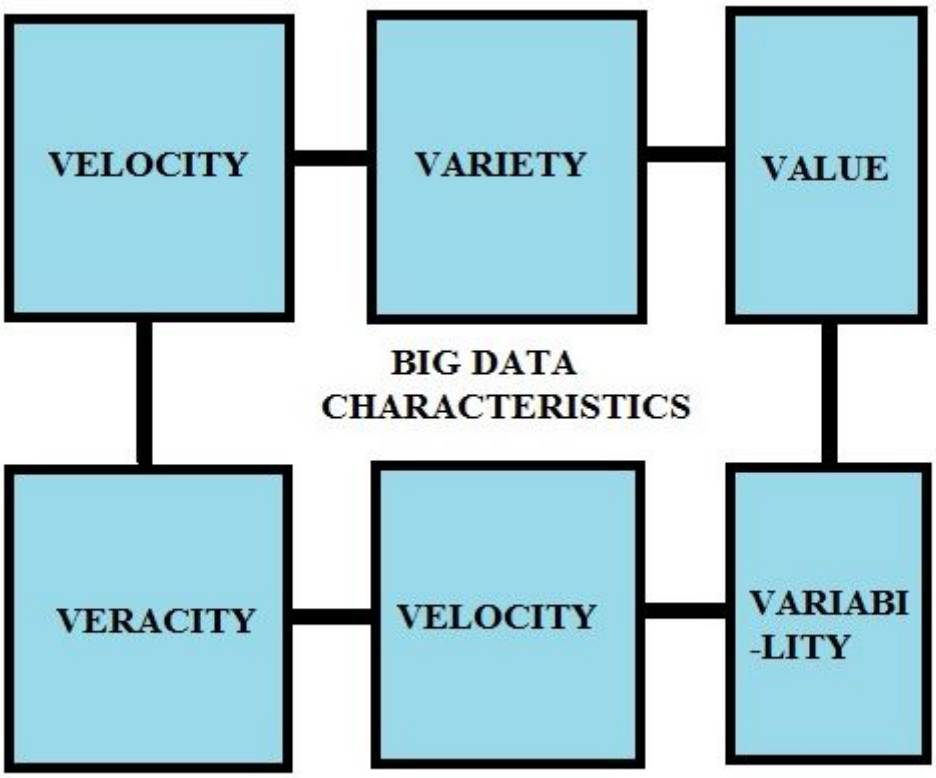
1) It is difficult to analyse and aggregate unstructured data such as test results, scanned documents, visual data and progress notes in patient electronic health record, etc.

2) The data in many health care providers are often segmented. Clinical data such as patient electronic health record consists of test results, images and progress notes. Quality and outcomes data such as patient's falls, surgical site infections, etc. are in risk or quality management department where it needs standards for validating, consolidating and processing data are needed. 3) Privacy issues in the patients data such as health records , insurance details,etc. even if the privacy of the patient is protected , many health care providers are unwilling to share data due to market competition 4) Collected data can be damaged or leaked through hackers.

12 VII. Discussion and Conclusion

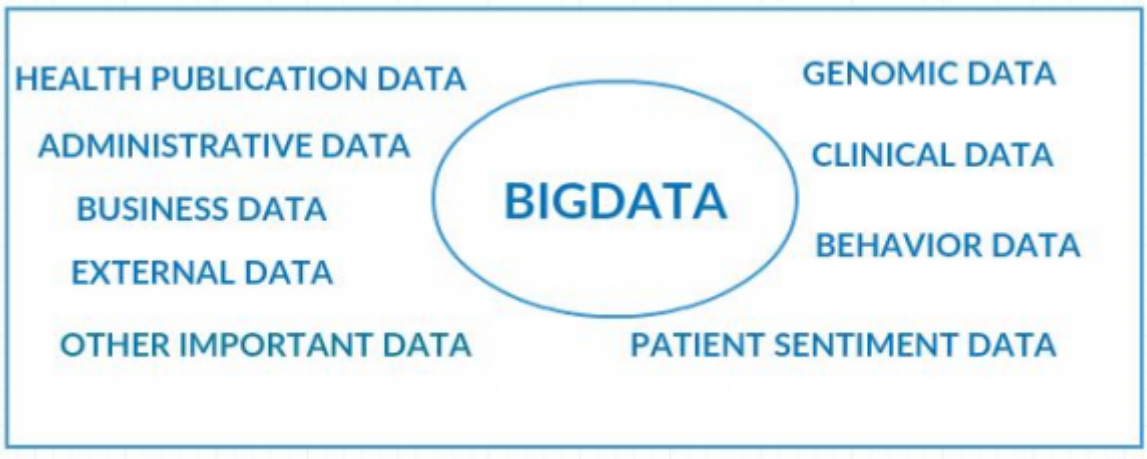
Big data is based on data generated from whole process of diagnosis and response of each case. It can lead to develop predictive models to determine which patients and health care users are mostly benefit from care management plan. By means of data analytics, it offers disease prevention, reduce in medical errors and better outcomes. it can improves and develops new therapies and research and development and has great potential to improve guide clinicians and medicines in delivering value based care. Big data has challenges in arena of bio-medical and health care to overcome such as information security, lack of infrastructure, data privacy and leakage, etc.

Utmost of all challenges can be scope for future research topics, the following topics may have a chance of future research: medical data confidentiality and interoperability, indexing and processing of continuous data, analyzing and aggregating of unstructured biomedical and health care data, security of health care data, etc.



1

Figure 1: Fig. 1 :



2

Figure 2: Fig. 2 :

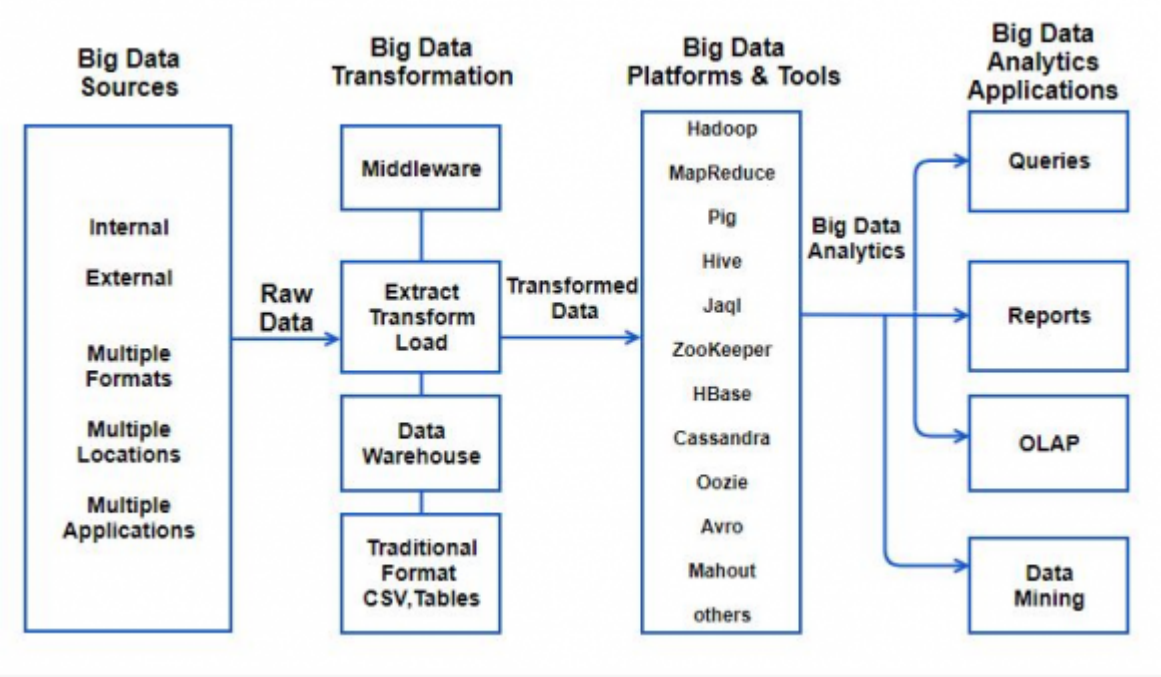


Figure 3:

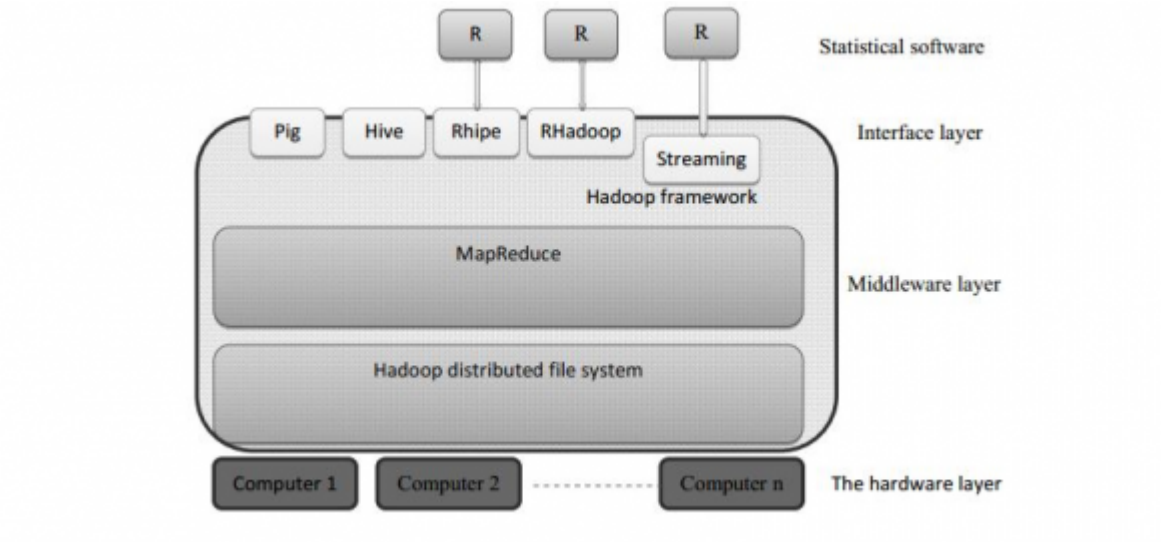
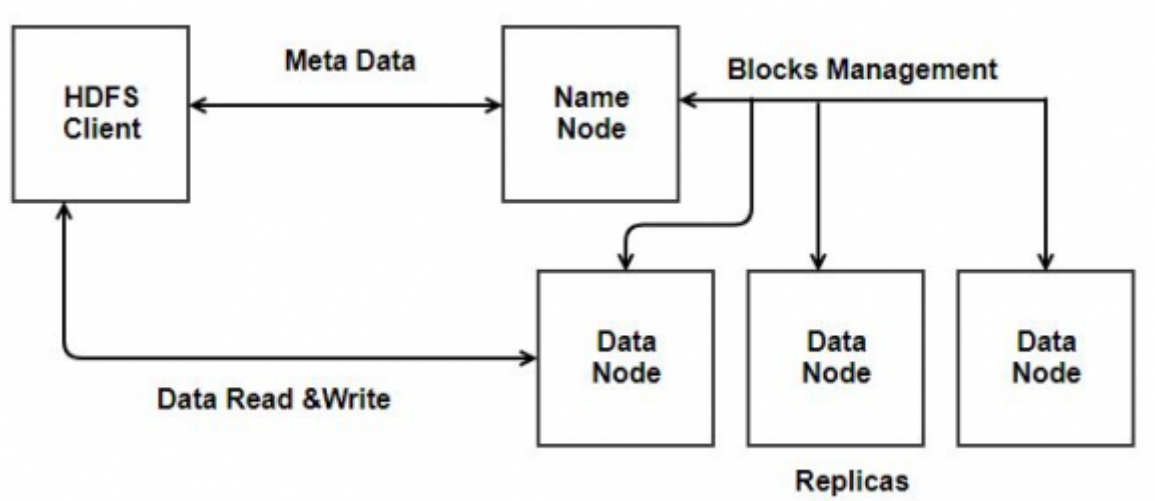
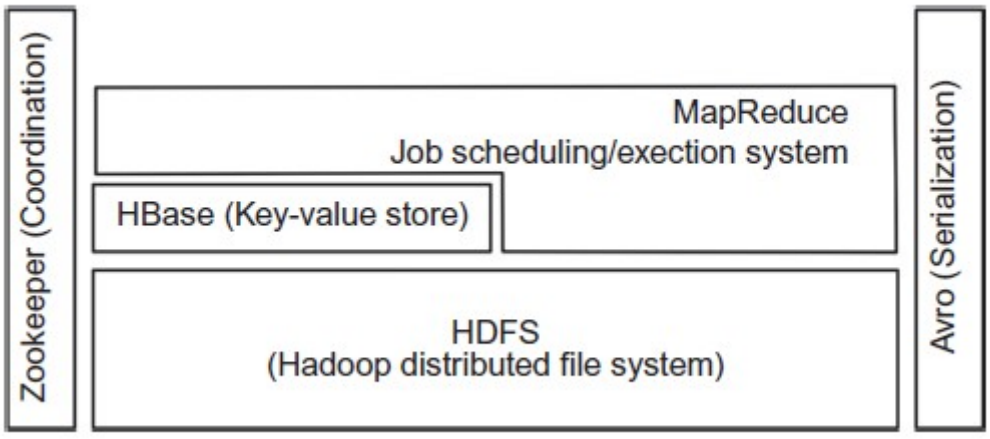


Figure 4:



3

Figure 5: Fig. 3 :



4

Figure 6: Fig. 4 :

this paper focused on study on potential of big data in biomedical and health care arena and their application, challenges and opportunities. [1] [2] [3]

[1] © 2017 Global Journals Inc. (US)

[2] © 2017 Global Journals Inc. (US)

[3] © 2017 Global Journals Inc. (US) Year 2017 Potential of Big Data Analytics in Bio-Medical and Health Care Arena: An Exploratory Study

[Turk ()] *A chart of the big data ecosystem*, M Turk . 2012.

[Ji ()] *Applications analysis of big data analysis in the medical industry, international journal of database theory and application*, Zhendong Ji . 2015. 8 p. .

[Raghupathi and Raghupathi ()] 'Big data analytics in healthcare: Promise and potential. Health Inform'. W Raghupathi , V Raghupathi . 10.1186/2047-2501-2-3. *Science Syst* 2014. 2 p. .

[Shrestha ()] 'Big data and cloud computing'. R B Shrestha . *Applied Radiology* 2014.

[Tene and Polonetsky ()] 'Big data for all: Privacy and user control in the age of analytics'. O Tene , J Polonetsky . *Northwestern J. Intellectual Property* 2013. 2017. 11 p. . (30 Year)

[Wang ()] 'Big Data in Medical Applications and Health Care'. Lidong Wang , CherylAnn , Alexander . *American Medical Journal* 2015. 6 (1) .

[Liang and Kelemen ()] 'Big Data Science and its Applications in Healthcare and Medical Research: Challenges and Opportunities'. Y Liang , A Kelemen . *Austin Biom and Biostat* 2016. 3 (1) .

[Zolfagha et al. ()] 'Big Data Solutions for Predicting Risk of Readmission for Congestive Heart Failure Patients'. Kiyana Zolfagha , Naren Meadem , Ankur Teredesai , Senjuti Basu Roy , Brian Muckian . *IEEE International Conference of big data*, 2013. p. .

[Sean et al. ()] *methods of using real-time social media technologies for detection and remote monitoring of HIV outcomes*, D Sean , Caitlin Young , Bryan Rivers , Lewis . 2014. 63 p. .

[Hsieh et al. ()] 'Mobile, cloud and big data computing: Contributions, challenges and new directions in tele cardiology'. J C Hsieh , A H Li , C C Yang . DOI: 10.3390/I jerph10116131. *Int. J. Environ.Res. Public Health* 2013. 10 p. .

[Nambisan et al. ()] 'Social media, big data and public health informatics: Ruminating behaviour of depression revealed through twitter'. Priya Nambisan , Zhihui Luo , Akshat Kapoor , Timothy B Patrick , Ron A Cisler . *th Hawaii International Conference on system sciences*, 2015. 48 p. .

[Zou et al. ()] 'Survey of Mapreduce frame operation in bioinformatics'. Quan Zou , Xu-Bin , Wen-Rui Li , Zi-Yu Jiang , Gui-Lin Lin , Ke Li , Chen . *briefings in bioinformatics* 2013. 15 (4) p. .

[Schultz ()] 'Turning healthcare challenges into big data opportunities: A use-case review across the pharmaceutical development lifecycle'. T Schultz . 10.1002/bult.2013.1720390508. *Bull. Association Inform. Sci. Technol* 2013. 39 p. .