

Big Data in Medical Applications and Health Care

¹Lidong Wang and ²Cheryl Ann Alexander

¹Department of Engineering Technology, Mississippi Valley State University, USA

²Department of Nursing, University of Phoenix, USA

Article history

Received: 31-3-2015

Revised: 30-4-2015

Accepted: 4-5-2015

Corresponding Author:

Lidong Wang

Department of Engineering
Technology, Mississippi Valley
State University, USA

Email: lwang22@students.tntech.edu

Abstract: Big Data can unify all patient related data to get a 360-degree view of the patient to analyze and predict outcomes. It can improve clinical practices, new drug development and health care financing process. It offers a lot of benefits such as early disease detection, fraud detection and better healthcare quality and efficiency. This paper introduces the Big Data concept and characteristics, health care data and some major issues of Big Data. These issues include Big Data benefits, its applications and opportunities in medical areas and health care. Methods and technology progress about Big Data are presented in this study. Big Data challenges in medical applications and health care are also discussed.

Keywords: Big Data, Health Care, Medicine, Diagnosis, Tele-diagnosis, Privacy, Monitoring, Clinical Practice, Personalized Patient Care, Fraud Detection, E-Consultation, Big Data Analytics, Data Mining, Machine Learning, Hadoop, Smart Health, Information Security

Introduction

Big Data Concept and Characteristics

Big data is the data that exceeds the processing capacity of conventional database systems. The data is too big, moves too fast, or doesn't fit the strictures of conventional database architectures (Dumbill, 2013). Big data characteristics can be described by "6Vs". They are: Volume, Velocity, Variety, Value, Variability and Veracity (Russom, 2011; Eaton *et al.*, 2012; O'Reilly Radar Team, 2012; Zikopoulos *et al.*, 2012; Bellini *et al.*, 2013; Demchenko *et al.*, 2013; Megahed and Jones-Farmer, 2013; Minelli *et al.*, 2013; Rajpathak and Narsingpurkar, 2013):

- **Volume:** This means data size such as Terabytes (TB: Approximately 10^{12} bytes), Petabytes (PB: Approximately 10^{15} bytes) and Zettabytes (ZB: Approximately 10^{21} bytes), etc
- **Velocity:** Data is generated at a high speed
- **Variety:** This represents all types of data such as structured data from relational tables, semi-structured data from key-value web clicks and unstructured data from email messages, articles and streamed video and audio, etc
- **Value:** It is defined by the added-value that the collected data can bring. It refers to the value that the data adds to creating knowledge. There is some valuable information somewhere within the data

- **Variability:** It refers to data changes during processing and lifecycle. Increasing variety and variability also increases the attractiveness of data and the potentiality in providing unexpected, hidden and valuable information
- **Veracity:** It includes two aspects: Data consistency (or certainty) and data trustworthiness. Data can be in doubt: incompleteness, ambiguities, deception and uncertainty due to data inconsistency, etc

There is often noisy data or false information in big data. The focus of Big Data is on correlations, not causality (Bottles and Begoli, 2014). In addition, the data we consider big today may not be considered big tomorrow because of the advances in data processing, storage and other system capabilities (Zaslavsky *et al.*, 2012).

Health Care Data

Big data in healthcare refers to electronic health data sets so large and complex that it is difficult to manage with traditional or common data management methods and traditional software and/or hardware (Priyanka and Kulennavar, 2014). Some health care data are characterized by a need for timeliness; for example, data generated by wearable or implantable biometric sensors; blood pressure, or heart rate is often required to be collected and analyzed in real-time (Helm-Murtagh, 2014).

Data in healthcare can be categorized as follows:

Genomic Data

It refers to genotyping, gene expression and DNA sequence (Chen *et al.*, 2012; Priyanka and Kulennavar, 2014).

Clinical Data and Clinical Notes

About 80% of this type data are unstructured documents, images and clinical or transcribed notes (Yang *et al.*, 2014):

- Structured data (e.g., laboratory data, structured EMR/HER)
- Unstructured data (e.g., post-op notes, diagnostic testing reports, patient discharge summaries, unstructured EMR/HER and medical images such as radiological images and X-ray images)
- Semi-structured data (e.g., copy-paste from other structure source)

Behavior Data and Patient Sentiment Data

- Web and social media data
Search engines, Internet consumer use and networking sites (Facebook, Twitter, LinkedIn, blog, health plan websites and smartphone, etc.) (Terry, 2013)
- Mobility sensor data or streamed data (data in motion, e.g., electroencephalography data)
They are from regular medical monitoring and home monitoring, telehealth, sensor-based wireless and smart devices (Shrestha, 2014)

Health Publication and Clinical Reference Data

Text-based publications (journals articles, clinical research and medical reference material) and clinical text-based reference practice guidelines and health product (e.g., drug information) data (Miller, 2012; Priyanka and Kulennavar, 2014).

Administrative, Business and External Data

- Insurance claims and related financial data, billing and scheduling (Terry, 2013)
- Biometric data: Fingerprints, handwriting and iris scans, etc

Other Important Data

- Device data, adverse events and patient feedback, etc. (Yang *et al.*, 2014)
- The content from portal or Personal Health Records (PHR) messaging (such as e-mails) between the patient and the provider team; the data generated in the PHR

Benefits of Big Data in Medical Applications and Health Care

Effective large-scale analysis often requires the collection of heterogeneous data from multiple sources. For example, obtaining the 360-degrees health view of a patient (or a population) benefits from integrating and analyzing the medical health record along with Internet available environmental data and then even with readings from multiple types of meters (for example, glucose meters, heart meters, accelerometers, among others) (Jagadish *et al.*, 2014).

Applying advanced analytics to patient profiles, characteristics and the cost and outcomes of care can help identify the most clinically and cost effective treatments, proactively identify individuals who would benefit from preventative care or lifestyle changes. Big Data could help reduce waste and inefficiency in the following three areas (Manyika *et al.*, 2011):

- *Clinical operations*: Determine more clinically relevant and cost-effective ways to diagnose and treat patients
- *Research & development*: (1) lower attrition and produce a leaner, faster, more targeted R&D pipeline in drugs and devices with the help of predictive modeling; (2) improve clinical trial design, thus reducing trial failures and speeding new treatments to market; and (3) identify follow-on indications and discover adverse effects before products reach the market
- *Public health*: (1) analyze disease patterns and track disease outbreaks and transmission to improve public health surveillance and speed response; (2) faster develop more accurately targeted vaccines and (3) turn large amounts of data into actionable information

Big Data benefits in medical applications and health care can be summarized as follows (Helm-Murtagh, 2014; Raghupathi and Raghupathi, 2014): (1) Improvement of health outcomes through more accurate and precise diagnoses; identification of patients who are at risk of adverse outcomes; and customization of care at the level of the individual patient (personalized medicine). (2) Reduction of costs through earlier detection of disease; elimination of unnecessary and duplicate care; reduction in variations in care; and elimination of erroneous and improper claims submissions. (3) Predicting and managing obesity and health risks; detecting health care fraud more quickly and efficiently (Certain developments or outcomes may be predicted and/or estimated based on vast amounts of historical data). (4) Decreasing inappropriate Emergency Department (ED) utilization by using statistical models to identify the best ED services or care alternatives that are more appropriate, more convenient and lower in cost according to health conditions, prior use of health care

resources (e.g., having a primary care provider) and proximity to sites of care. (5) Providing advantages to Health Informatics. This is fulfilled by allowing for more tests cases or more features for research, leading to both quicker validation of studies and the ability to accrue enough instances for training. Big Data approaches have been used for the analysis of Health Informatics data gathered at multiple levels, including the molecular, tissue, patient and population levels. The amount of data produced within Health Informatics has grown to be quite vast. Big Data analytics grants potentially great possibilities to gain much knowledge in Health Informatics.

Applications and Opportunities of Big Data in Medical Applications and Health Care

Big Data can provide support across all aspects of health care. Big Data analytics has gained traction in genomics, epidemic spread prediction, clinical outcome, fraud detection, pharmaceutical development and personalized patient care, etc. There are potential applications in these areas. The specific applications of Big Data in the areas are as follows.

Genomics Analytics

Genomic data is becoming critical to the complete patient record. Combining patient genomic data with clinical data helps cancer treatment (Chen *et al.*, 2012; Priyanka and Kulennavar, 2014).

Flu Outbreak Prediction and Control

In public and population health, continuously aggregating and analyzing public health data helps detect and manage potential disease outbreaks. Big Data analytics can mine web-based and social media data to predict flu outbreaks based on consumer search, social content and query activity (Priyanka and Kulennavar, 2014).

Clinical Outcome Analytics

Clinical analytics can be performed through unifying clinical, financial and operation data for efficient clinical decisions. Blue Cross and Blue Shield of North Carolina, USA has provided several promising examples of how Big Data can be used to reduce the cost of care, predict and manage health risks and improve clinical outcomes (Helm-Murtagh, 2014).

Fraud Detection and Prevention

Identifying, predicting and minimizing fraud can be implemented by using advanced analytic systems for fraud detection and checking the accuracy and consistency of claims. Big Data predictive modeling can be used by health care payers for fraud prevention. Fraud waste and abuse analytics can be performed in analyzing claims and benefits of Veterans benefits and education fraud (White, 2014; Raghupathi and Raghupathi, 2014).

Medical Device Design and Manufacturing

Big Data tools enable a broader set of anatomical configurations, device materials, delivery methods and tissue interactions to be evaluated. Computational methods and Big Data can play an important role in medical device design and manufacturing (Erdman and Keefe, 2013).

Personalized Patient Care

Healthcare is moving from a disease-centered model towards a patient-centered model. In a disease-centered model, physicians' decision making is centered on the clinical expertise and data from medical evidence and various tests. In a patient-centered model, patients actively participate in their own care and receive services focused on individual needs and preferences. The patient-centric model creates a personalized disease risk profile, as well as a disease management plan and wellness plan for an individual. Personalized healthcare is a data-driven approach. With the increase in the use of electronic medical records, Big Data will facilitate to bring proactive and personalized patient care (Chawla Davis, 2013). In the near future, new big data-derived linkages will prompt timely updates of patient triage, diagnostic assistance and clinical guidelines to allow more precise and personalized treatment to improve clinical outcome for patients (Yang *et al.*, 2014).

E-Consultation and Tele-Diagnosis

In the future, the aggregated ECG and images from hospitals worldwide will become big data, which should be used to develop an e-consultation program helping on-site practitioners deliver appropriate treatment. Real-time tele-consultation and tele-diagnosis of ECG and images can be practiced via an e-platform for clinical, research and educational purposes. Big Data analytics can predict over 50% deaths with fewer false positives as compared with the traditional ECG analysis, conducted based on a smaller segment of ECG signals (Hsieh *et al.*, 2013).

Pharmaceuticals and Medicine

The ability of pharmaceutical companies to continue bringing new life-saving/life enhancing medicines to patients in a timely, yet cost-effective manner will depend on their ability to manage big data generated during all phases of pharmaceutical development. Integration of clinical, healthcare, patents, safety and public research data will provide key insights into decision making for target selection and lead optimization through Big Data analytics for drug discovery (Schultz, 2013).

Medical Education

Visual analytics was explored as a tool for finding ways of representing big data from the medical curriculum of an undergraduate medical program.

Table 1. BI&A and Big Data in smart health and well being

Applications	<ul style="list-style-type: none"> • Human and plant genomics • Health care decision support • Patient community analysis
Data	<ul style="list-style-type: none"> • Genomics and sequence data • Electronic Health Records (EHR), Personal Health Records (PHR) Health and patient social media <p><i>Characteristics:</i> Disparate but highly linked content, person-specific content, HIPAA, IRB and ethics issues</p>
Analytics	<ul style="list-style-type: none"> • Genomics and sequence analysis and visualization • EHR association mining and clustering • Health social media monitoring and analysis • Health text analytics • Health ontologies • Patient network analysis • Adverse drug side-effect analysis • Privacy-preserving data mining Impacts Improved health care quality, improved long-term care, patient empowerment
Impacts	Improved healthcare quality, improved long-term care, patient empowerment.

IRB: Institutional Review Board

HIPAA: Health Insurance Portability and Accountability Act

A possible use of Big Data in the medical education context (Vaitsis *et al.*, 2014) is to: (1) identify data connections and the relations between them; (2) determine data's roles in the lowest level of a course and in the overall picture of the medical program; (3) perceive and analyze the curriculum in terms of identifying whether knowledge, skills and attitude are constructed through the alignment of teaching methods and assessment towards learning outcomes and (4) perform gap analysis by comparing different states in which data can be found to identify possible discrepancies.

Smart Health and Wellbeing

Business Intelligence and Analytics (BI&A) and the related field of Big Data analytics have become increasingly important in the business communities. Table 1 (Chen *et al.*, 2012) summarizes some BI&A features and capabilities in smart health and wellbeing, including applications, data characteristics, analytics and potential impacts.

Big Data has brought great opportunities in medical applications and health care. Big Data applications will expand to more areas (such as telemedicine and digital hospitals), further improve medical service and deliver extensive value-based care. Big Data applications and opportunities need technology support.

Methods and Technology Progress in Big Data

In healthcare/medical field, large amount of information about patients' medical histories, symptomatology, diagnoses and responses to treatments and therapies is collected. Data mining techniques can be implemented to derive knowledge from this data in order to either identify new interesting patterns in infection control data or to examine reporting practices. Moreover, predictive models can be used as detection tools exploiting Electronic Patient Record (EPR) accumulated for each person of the area (Bellini *et al.*, 2013).

For Big Data healthcare systems, the Hadoop-MapReduce framework is uniquely capable of storing a wide range of healthcare data types including electronic medical records, genomic data, financial and claims data etc. and offers high scalability, reliability and availability than traditional Database Management Systems (DBMS). In addition, intelligent functional modules such as specialized machine-learning algorithms for image analysis and recognition, diagnosis, surveillance, detection, notification etc., can be built on it (Ngufor and Wojtusiak, 2013).

Figure 1 shows a general framework of big data and big data analytics.

In order to create an automatic lesion diagnostic model, an automatic breast cancer diagnostic model was created using a pattern recognition algorithm and big data mining technique. Data mining is the process of discovering useful correlations hidden in large quantities data and extracting information which can be used in decision-making. The Support Vector Machine (SVM) algorithm that is most frequently used with an elevated accuracy was also presented in detail. A machine learning algorithm inputs data into the computer establishes criteria for categorization and predicts the category of the data when new data are input (Lee and Lee, 2014).

Visual analytics presents an area of synergistic research with big data by conceptualizing the output of complex processes through intuitive graphical means. Metrics dash boarding, real-time interactive visualization and Giga-node graph exploration are some examples that would serve as appropriate visualization solutions to the big data examples. Unstructured data needs to be converted into analysis-ready datasets, which include comprehensive workflows for of big data solutions. Consideration of the structure of the end-data models is vital for the visualization process (Schultz, 2013).

Visual analytics combines data analysis and manipulation techniques, information and knowledge representation and human cognitive strength to perceive and recognize visual patterns (Vaitsis *et al.*, 2014). Visual Analytics supports Big Data by providing interactive visualizations that allow people to navigate these datasets. Visual Analytics has been defined as "the science of analytical reasoning facilitated by interactive visual interfaces" (Thomas and Cook, 2006).

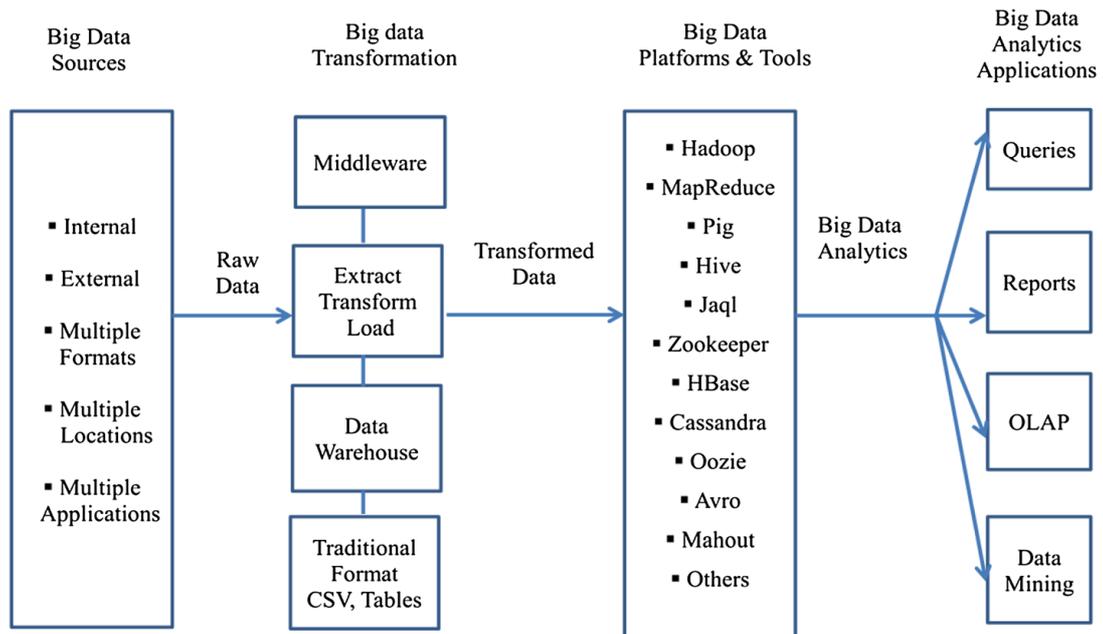


Fig. 1. An applied conceptual architecture of Big Data analytics (Raghupathi and Raghupathi, 2014) (OLAP: Online Analytical Processing; CSV: Comma Separated Values)

Big Data enabled by cloud technologies could provide us new insights-clinically, operationally and in research (Shrestha, 2014). The concept of storage-as-a-service cloud computing, which provides hospitals with a big data storage capacity based on their specific demands at a low cost. In cardiology, cloud computing technology and mobile teleconsultation should be combined because mobile teleconsultation requires high speed data delivery and a big data center where data can be delivered, stored, retrieved and managed securely (Hsieh *et al.*, 2013).

Besides general cloud infrastructure services (storage, compute, infrastructure/VM management), the following services are required to support Big Data (Turk, 2012):

- Cluster services
- Hadoop related services and tools
- Specialist data analytics tools (logs, events, data mining, etc.)
- Databases/Servers SQL, NoSQL
- MPP (Massively Parallel Processing) databases
- Registries, indexing/search, semantics, namespaces
- Security infrastructure (access control, policy enforcement, confidentiality, trust, availability, privacy)

Organizations used various methods of de-identification (anonymization, pseudonymization, encryption, key-coding, data sharing) to distance data from personal identities and preserve individuals' privacy. De-identification has been viewed as an important protective measure to be taken under the data

security and accountability principles. Yet, over the past few years, computer scientists have repeatedly shown that even anonymized data can typically be re-identified and associated with specific individuals. De-identified data, in other words, is a temporary state rather than a stable category (Tene and Polonetsky, 2013).

Challenges of Big Data in Medical Applications and Health Care

Large volume, velocity and variety of big data have brought big challenges in data storage, curation, retrieval, search and visualization. Variability and veracity of big data indicate data instability and uncertainty, which often makes Big Data analytics difficult.

Major challenges of Big Data in medical applications and healthcare are as follows: (1) the data in many health care providers, specifically hospitals, are often segmented or siloed. Clinical data such as patient history, vital signs, progress notes and diagnostic test results are stored in the EHR. Quality and outcomes data such as surgical site infections, rates of return to surgery and patient falls are in the quality or risk management departments. Standards for validating, consolidating and processing data are needed (White, 2014). (2) It is difficult to aggregate and analyze unstructured data. Unstructured data include: Test results, scanned documents, images and progress notes in the patients' EHR, etc. (White, 2014). Efficiently handling large volumes of medical imaging data, extracting potentially useful information and biomarkers and understanding unstructured clinical notes in the right context are

challenges (Priyanka and Kulennavar, 2014). (3) Analyzing genomic data is a computationally intensive task; combining with standard clinical data adds additional layers of complexity (Priyanka and Kulennavar, 2014). (4) An emerging new data source is telemetry from patient-owned devices and information entered by patients. The challenge of Big Data becomes even greater when telemetry from automated monitoring devices is included. Such data could include subjective symptom scores (pain, mood and mobility); patient reported outcomes; and device telemetry such as weight, activity, glucose, blood pressure and pulse oximetry (Halamka, 2014). The capture, indexing and processing of continuously streaming (and possibly annotated) fine-grained temporal data is a challenge (Schultz, 2013). (5) Big Data's focus on correlations, not causality, is difficult for physicians biased toward the biomedical model, where the focus is finding the cause of the disease in order to effectively treat it. Big data means more information, but there is often noisy data or false information (Bottles and Begoli, 2014). (6) Privacy issues in the Health Insurance Portability and Accountability Act (HIPAA) are often cited as barriers to collecting big data (Warner, 2013). In telecardiology and tele-consultation, data confidentiality in the cloud, data interoperability among hospitals and network latency and accessibility are challenges (Hsieh *et al.*, 2013). (7) Even if the privacy of the patient can be protected, many health care providers are reluctant to share data because of market competition. It is difficult to determine the proper balance between protecting the patient's information and maintaining the integrity and usability of the data. Open access, integration, standardization of readable and useable data is a challenge (White, 2014). (8) Data hackers have become more damaging in big data. Data leakage can be costly. In March 2012, hackers broke into Utah's Department of Health database and downloaded personal data from 780,000 patients (Social Security Numbers were downloaded for 280,000 patients) (Schmitt *et al.*, 2013). Biometrics such as a fingerprint helps improve information security and protect against data leakage. However, it is almost impossible to guarantee complete data security. (9) Both providers and payers pointed to resource shortfalls such as staffing, budget and infrastructure as the big barriers to the adoption of Big Data. Lack of infrastructure and policies, standards and practices that make the most of big data in healthcare were also cited as a concern (Bulletin Board, 2014). (10) De-identification is the process by which personally identifiable information is removed from health information so that there cannot be any linkage back to the individual in any way. HIPAA outlines two procedures for de-identifying the information: Safe harbor and expert determination. The ability to gather and analyze de-identified data is essential to driving down cost and improving quality. Concerns exist that data cannot really be fully de-identified (Warner, 2013).

Big Data technology challenges such as data integration, data visualization and information security will be overcome with the advances of computer science, scientific computation and other disciplines. Other challenges such as standards, data privacy and ownership and data sharing and cross-disciplinary collaboration, etc. need supports from agencies and governments in policies. It is important and necessary to consolidate e-Infrastructures as persistent platforms to ensure continuity in Big Data.

Conclusion and Future Research

Big Data is based on data obtained from the whole process of diagnosis and treatment of each case. Big Data analytics can perform predictive modeling to determine which patients are most likely to benefit from a care management plan. It is moving forward quickly in population health and quality measurement. Big Data offers a lot of benefits such as disease prevention, reduced medical errors and the right care at the right time and better medical outcomes. In addition, Big Data can improve the Research and Development (R&D) and translation of new therapies. Big data has great potential to improve medicine, guide clinicians in delivering value-based care.

Big Data has challenges in medical applications and healthcare. These challenges include consolidating and processing segmented or siloed data, aggregating and analyzing unstructured data, indexing and processing continuously streaming data, privacy, data leakage, information security and lack of infrastructure and unified standards, etc.

Most of the above challenges can be future research topics. These future research topics can be: Aggregating and analyzing unstructured health care data, indexing and processing of continuously stream data, medical data confidentiality and interoperability, health care data security and e-Infrastructures as persistent platforms for health care big data, etc. The authors of the paper will focus on Big Data in medical sensor data and streaming data processing, privacy-preserving data mining in health care, sentiment analysis of medical big data and personalization and behavioral modeling.

Acknowledgment

This study was supported in part by Technology and Healthcare Solutions, Inc. in Mississippi, USA. No conflict of interest to disclose.

Funding Information

The authors have no funding to report.

Author's Contributions

Lidong Wang and Cheryl Ann Alexander worked on the manuscript as group efforts at all stages.

Ethics

This article is original and contains unpublished material. The corresponding author confirms that both the authors have read and approved the manuscript and no ethical issues involved.

References

- Bellini, P., M. d. Claudio, P. Nesi and N. Rauch, 2013. Taxonomy and Review of Big Data Solutions Navigation. In: Big Data Computing, Akerkar, R. (Ed.), Chapman and Hall/CRC, pp: 57-101.
- Bottles, K. and E. Begoli, 2014. Understanding the pros and cons of big data analytics. *Physician Exec.*, 40: 6-12.
- Bulletin Board, 2014. Industry assesses potential, challenges of big data. *J. AHIMA*.
- Chawla, N.V. and D.A. Davis, 2013. Bringing big data to personalized healthcare: A patient-centered framework. *J. General Internal Med.*, 28: S660-S665. DOI: 10.1007/s11606-013-2455-8
- Chen, H.C., R.H.L. Chiang and V.C. Storey, 2012. Business intelligence and analytics: From big data to big impact. *MIS Q.*, 36: 1165-1188.
- Demchenko, Y., P. Grosso and C. de Laat, 2013. Addressing big data issues in scientific data infrastructure. Proceedings of the International Conference on Collaboration Technologies and Systems, May 20-24, IEEE Xplore Press, San Diego, CA, USA, pp: 48-55. DOI: 10.1109/CTS.2013.6567203
- Dumbill, E., 2013. Making sense of Big Data. *Big Data*.
- Eaton, C., D. Deroos, T. Deutsch, G. Lapis and P. Zikopoulos, 2012. Understanding big data. McGraw-Hill Companies.
- Erdman, A.G. and D.F. Keefe, 2013. Grand challenge: Applying regulatory science and big data to improve medical device innovation. *IEEE Trans. Biomed. Eng.*, 60: 700-706. DOI: 10.1109/TBME.2013.2244600
- Halamka, J.D., 2014. Early experiences with big data at an academic medical center. *Health Affairs*, 33: 1132-1138. DOI: 10.1377/hlthaff.2014.0031
- Helm-Murtagh, S.C., 2014. Use of Big Data by Blue Cross and Blue Shield of North Carolina. *NCMJ*, 75: 195-197.
- Hsieh, J.C., A.H. Li and C.C. Yang, 2013. Mobile, cloud and big data computing: Contributions, challenges and new directions in telecardiology. *Int. J. Environ. Res. Public Health*, 10: 6131-6153. DOI: 10.3390/ijerph10116131
- Jagadish, H.V., J. Gehrke, A. Labrinidis, Y. Papakonstantinou and J.M. Patel *et al.*, 2014. Big data and its technical challenges. *Communications ACM*, 57: 86-94. DOI: 10.1145/2611567
- Lee, J.H. and E.J. Lee, 2014. Computer-aided diagnosis sensor and system of breast sonography: A clinical study. *Sensors Transducers*, 180: 1-10.
- Manyika, J., M. Chui, B. Brown, J. Buhin and R. Dobbs *et al.*, 2011. Big data: The next frontier for innovation, competition and productivity. USA: McKinsey Global Institute.
- Megahed, F.M. and L.A. Jones-Farmer, 2013. A Statistical Process Monitoring Perspective on "Big Data", *Frontiers in Statistical Quality Control*, 11th Edn., Springer, New York.
- Miller, K., 2012. Big data analytics in biomedical research. *Biomedical Computation Review*.
- Minelli, M., M. Chambers and A. Dhiraj, 2013. Big Data, Big Analytics: Emerging Business Intelligence and Analytic Trends for Today's Business. 1st Edn., John Wiley & Sons, Inc. Hoboken, New Jersey, USA, ISBN-10: 111814760X, pp: 187.
- Ngufor, C. and J. Wojtusiak, 2013. Learning from large-scale distributed health data: An approximate logistic regression approach. Proceedings of the 30th International Conference on Machine Learning, Atlanta, Georgia, USA, *JMLR: W&CP*, pp: 1-8.
- O'Reilly Radar Team, 2012. Planning for big data. O'Reilly.
- Priyanka, K. and N. Kulennavar, 2014. A survey on big data analytics in health care. *Int. J. Comput. Sci. Inform. Technologies*, 5: 5865-5868.
- Raghupathi, W. and V. Raghupathi, 2014. Big data analytics in healthcare: Promise and potential. *Health Inform. Science Syst.*, 2: 1-10. DOI: 10.1186/2047-2501-2-3
- Rajpathak, T. and A. Narsingpurkar, 2013. Managing knowledge from big data analytics in product development. White Paper, Tata Consultancy Services.
- Russom, P., 2011. Big data analytics. TDWI Best Practices Report, Fourth Quarter.
- Schmitt, C., M. Shoffner and P. Owen, 2013. Security and privacy in the era of big data: The SMW, a technological solution to the challenge of data leakage. RENCI, University of North Carolina at Chapel Hill.
- Schultz, T., 2013. Turning healthcare challenges into big data opportunities: A use-case review across the pharmaceutical development lifecycle. *Bull. Association Inform. Sci. Technol.*, 39: 34-40. DOI: 10.1002/bult.2013.1720390508
- Shrestha, R.B., 2014. Big data and cloud computing. *Applied Radiology*.
- Tene, O. and J. Polonetsky, 2013. Big data for all: Privacy and user control in the age of analytics. *Northwestern J. Intellectual Property*, 11: 239-272.
- Terry, N.P., 2013. Protecting patient privacy in the age of big data. *UMKC Law Rev.*, 81: 385-415.

- Thomas, J.J. and K.A. Cook, 2006. A visual analytics Agenda. *IEEE Comput. Graphics Applic.*, 26: 10-13. DOI: 10.1109/MCG.2006.5
- Turk, M., 2012. A chart of the big data ecosystem, take 2.
- Vaitsis, C., G. Nilsson and N. Zary, 2014. Visual analytics in healthcare education: Exploring novel ways to analyze and represent big data in undergraduate medical education. *Peer J.*, 2: e683. DOI: 10.7717/peerj.683.
- Warner, D., 2013. Safe de-identification of big data is critical to health care. *Health Inform. Manage.*
- White, S.E., 2014. A review of big data in health care: Challenges and opportunities. *Open Access Bioinform.*, 6: 13-18. DOI: 10.2147/OAB.S50519
- Yang, S., M. Njoku and C.F. Mackenzie, 2014. 'Big data' approaches to trauma outcome prediction and autonomous resuscitation. *Brit. J. Hospital Med.*, 75: 637-641. DOI: 10.12968/hmed.2014.75.11.637
- Zaslavsky, A., C. Perera and D. Georgakopoulos, 2012. Sensing as a service and big data. *Proceedings of the International Conference on Advances in Cloud Computing (ACC' 12)*, Bangalore, India, pp: 1-8.
- Zikopoulos, P., C. Eaton, D. de Roos, 2012. *Understanding big data: Analytics for enterprise class hadoop and streaming data*. McGraw-Hill, New York.