Paradigm Shift of Statistical Big Data in Healthcare: Management, Analysis and Future Prospects

Dila Ram Bhandari
Lecturer, Nepal Commerce Campus, Tribhuvan University
Ron Kenett
Chairman, KPA Group
Professor, University of Turin in Italy and Senior Research Fellow,
Samuel Neaman Institute, Technion, Israel.

Abstract

'Big data' is vast describing volumes of information that can work miracles. Multiple attempts designed at highlighting the relationship between big data analytics and benefits for healthcare organizations have been raised. The big data influence on health organization management is still not clear due to the relationship's multi-disciplinary nature. It has become a topic of special interest for the past two decades because of a great potential that is hidden in it. Numerous public and private sector industries generate, store, and analyze big data with an aim to improve the services they provide. In the healthcare industry, various sources for big data include hospital records, medical records of patients, results of medical checkups, and devices that are a part of internet of things. Biomedical research also produces a significant share of big data pertinent to public healthcare. Big Data and healthcare are precarious to discourse the risk of hospitalization. This data requires proper management and analysis in order to derive expressive communication. There are various encounters associated with each step of handling big data which can only be surpassed by using high-end computing solutions for big data analysis. Providing relevant results for improving public health, healthcare providers are required to be fully equipped with appropriate infrastructure to systematically generate and analyze big data. A well-organized management, analysis, and interpretation of big data can transformation the game by opening new avenues for contemporary healthcare. With a strong integration of biomedical and healthcare data, modern healthcare institutes can probably revolutionize the medical therapies and personalized medicine.

Keywords: *Big data analytics, Healthcare, Biomedical research, Internet of things, Quantum computing, Information quality.*

Cite this paper
Bhandari, D. R. & Keneet, R. (2022). Paradigm Shift of Statistical Big Data in Healthcare: Management, Analysis and Future Prospects. *Pravaha, 28*(1), 35-44.

Introduction

Data have become an omnipresent concept in our daily lives with the routine collection, storage, processing and analysis of immense amount of data. Big Data is considered the most valuable and powerful fuel that can run the massive IT industries of the 21st Century. This specific is cross-sectorial, extending from the domain of machine learning and engineering, to finances and medicine. Big data is transforming and will transform the healthcare institutes in the nearby upcoming. Scientific literature in the managerial context applied to healthcare organizations, consider the Big Data Analytics (BDA) a fundamental tool, so much so that it has attracted the attention of the scientific community and stakeholders. The data has changed the way we manage, analyze, and leverage data across the most notable areas where data analytics is making big changes is healthcare. Information has been the key to a better institute and novel developments. Big data in healthcare is a term used to describe massive volumes of information created by the adoption of digital technologies that collect patients' records and help in managing hospital performance, else too large and complicated for usual technologies. The more information we have, the more optimally we can establish ourselves to provide the best results. So, data collection is an important part for every corporation. We can also use the data for the estimate of existing trends of evident parameters and future actions. Today, we are facing a situation wherein we are flooded with tons of data from every aspect of our life such as social activities, transportation, weather, science, work, health, etc. The technological improvements have helped us in creating more and more data, even to a level where it has become unmanageable with presently existing technologies. This has led to the formation of the term 'big data' to explain data that is large and riotous. In order to meet our present and future social essentials, we need to advance new policies to systematize this data and derive significant evidence. Unique social need is healthcare like every other industry, healthcare institutes are producing data at a tremendous rate that presents many benefits and confronts at the similar time.

The data overload

Every day, people working with various organizations around the world are producing a massive volume of data. The term "digital universe" defines such huge volumes of data created, replicated, and consumed in a specific year. International Data Corporation (IDC) estimated the approximate size of the digital universe in 2005 to be 130 exabytes (EB). The digital space in 2017 prolonged to about 16,000 EB. IDC predicted that the digital universe would expand to 40,000 EB by the year 2020. To imagine this size, we would have to assign about 5200 gigabytes (GB) of data to all individuals which explains the remarkable speed at which the digital universe is growing. The internet giants, like Google and Facebook, have been collecting and storing massive volumes of data. Google may store a variety of information including user location, advertisement preferences, list of applications used, internet browsing history, contacts, bookmarks, emails, and other necessary information accompanying with the user. Similarly, Facebook stores and analyzes more than about 30 petabytes (PB) of user-generated data.

Big Data Parameters

Big data management processes with the collection and retrieval of multiple, confusing, dynamic and large datasets. Given the fact that big data is unmanageable using the traditional software, we need technically advanced applications and software that can utilize fast and cost-efficient high-end computational power for such tasks. Implementation of artificial intelligence (AI) algorithms and novel fusion algorithms would be necessary to make sense from this large amount of data. With proper storage and analytical tools in hand, the information and insights derived from big data can make the critical social infrastructure components and services more aware, interactive and efficient. In addition, visualization of big data in a user-friendly manner will be a critical factor for societal development. Big data is typically defined by five 'Vs':

- Value: Valuable, reliable and truthful data that needs to be stored, processed, analyzed to find insights.
- Velocity: Big Data is to provide data on demand and at a speedier pace.
- Variety: Foremost part of the data is unstructured and irrelevant,
- Volume: Information generated every second from social media, cell phones, cars, credit cards, M2M sensors, images, video, and whatnot etc.
- Variability: Interacts with inconsistent data.

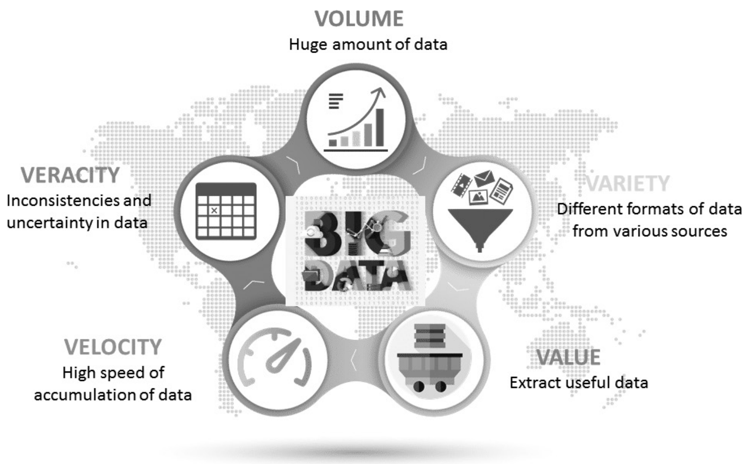

Fig. 1: Characteristics of big data

Healthcare as a big-data repository

Healthcare is a multi-dimensional practice known with the sole goal for the prevention, diagnosis, and treatment of health-related issues. The foremost mechanisms of a healthcare system are the health specialists, health amenities, and a financing institution. The health specialists belong to countless health divisions like dentistry, medicine, midwifery, nursing, psychology, rehabilitation, and many others. Healthcare is vital at numerous levels provisional on the urgency of situation. Specialists serve it as the first point of consultation (for primary care), serious care requiring skilled specialists (secondary care), advanced medical investigation and treatment (tertiary care) and highly unusual diagnostic and surgical procedures (quaternary care). At all these levels, the health professionals are responsible for different kinds of information such as patient's medical history, medical and clinical data (like data from imaging and laboratory examinations), and other personal medical data. The common practice to store such medical records for a patient was in the form of either handwritten minutes or typed details. The clinical illustration records freeze the episode of infection as a story in which patient, family and the doctor are a part of the plot". With the advent of computer systems and its potential, the digitization of all clinical exams and medical records in the healthcare systems has become a standard and widely adopted practice nowadays.

Digitization of healthcare and big data

In today's digital world, every individual seems to be obsessed to track their fitness and health statistics using the in-built pedometer of their portable and wearable devices such as, smartphones, smartwatches, fitness dashboards or tablets. With an increasingly mobile society in almost all aspects of life, the healthcare infrastructure needs remodeling to accommodate mobile devices. The practice of medicine and public health using mobile devices, known as mHealth or mobile health, pervades different degrees of health care especially for chronic diseases, such as diabetes and cancer. In fact, Apple and Google have developed devoted platforms like Apple's ResearchKit and Google Fit for developing research applications for fitness and health statistics. These applications support seamless interaction with various consumer devices and embedded sensors for data integration. These apps help the doctors to have direct access to your overall health data. Both the user and their doctors get to know the real-time status of your body.

EHRs have introduced many advantages for handling modern healthcare related data. The first advantage of EHRs is that healthcare professionals have an improved access to the entire medical history of a patient which includes medical diagnoses, prescriptions, data related to known allergies, demographics, clinical narratives, and the results obtained from various laboratory tests. The recognition and treatment of medical conditions thus is time efficient due to a reduction in the lag time of previous test results. Similar to EHR, an electronic medical record (EMR) stores the standard medical and clinical data gathered from the patients. EHRs, EMRs, individual health record, medical practice management software (MPM), and various other healthcare data components collectively have the potential to improve the quality, service efficiency, and costs of healthcare along with the reduction of medical errors. The big data in healthcare includes the healthcare payer-provider data along with the genomics-driven experiments (genotyping, gene expression data) and other data acquired from the smart web of internet of things (IoT). The management and usage of such healthcare data has been increasingly dependent on information technology and the development and usage of wellness monitoring devices and related software that can generate alerts and share the health-related data of a patient with the respective health care providers has gained momentum, especially in establishing a real-time biomedical and health monitoring system.

Healthcare industry has not been quick enough to adapt to the big data movement compared to other industries. One such source of clinical data in healthcare is 'internet of things' (IoT). The analysis of data collected from these chips or sensors may reveal critical information that might be beneficial in improving lifestyle, establishing measures for energy conservation, improving transportation, and healthcare. Such resources can interconnect various devices to provide a reliable, effective and smart healthcare service to the elderly and patients with a chronic illness. Using the web of IoT devices, doctors can measure and monitor various parameters from their clients in their respective locations like home or office. IoT devices used in healthcare include fitness or health-tracking wearable devices, biosensors, clinical devices for monitoring vital signs, and others types of devices or clinical instruments. Such IoT devices generate a large amount of health-related data.

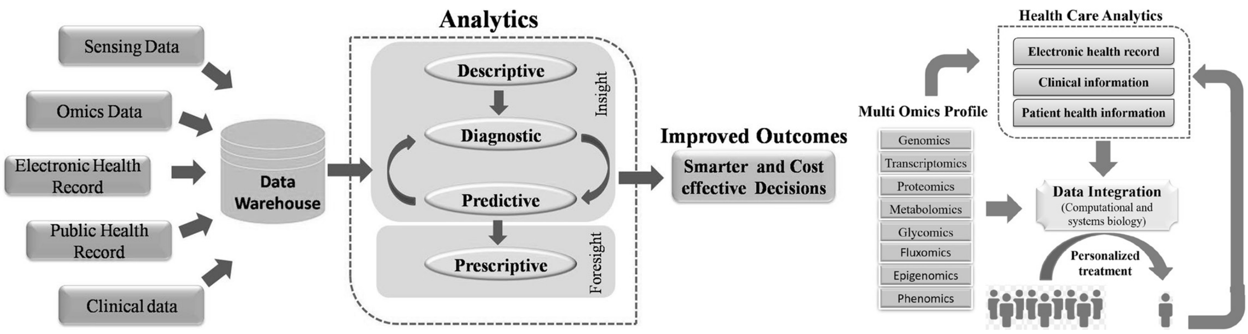

Fig.2: Health care Analytics

Management and analysis of big data

All in all, we've noticed three key trends of healthcare analytics: the patient experience will improve dramatically, including quality of treatment and satisfaction levels; the overall health of the population can also be enhanced on a sustainable basis, and operational costs can be reduced significantly. Big Data and predictive analytics can contribute to precision public health by improving public health surveillance and assessment, therefore, in a public health perspective, the gathering of a very large amount of data, constitute an inestimable resource to be used in epidemiological research, analysis of the health needs of the population, evaluation of population-based intervention and informed policy making. Big data is the huge amounts of a variety of data generated at a rapid rate in health care sectors. The data gathered from various sources is mostly required for optimizing consumer services rather than consumer consumption. To make it available for scientific community, the data is required to be stored in a file format that is easily accessible and readable for an efficient analysis. In the context of healthcare data, another major challenge is the implementation of high-end computing tools, protocols and high-end hardware in the clinical setting. It would enhance the efficiency of acquiring, storing, analyzing, and visualization of big data from healthcare. The main task is to annotate, integrate, and present this complex data in an appropriate manner for a better understanding.

Heterogeneity of data is another challenge in big data analysis. The huge size and highly heterogeneous nature of big data in healthcare renders it relatively less informative using the conventional technologies. Cloud computing is such a system that has virtualized storage technologies and provides reliable services. It offers high reliability, scalability and autonomy along with ubiquitous access, dynamic resource discovery and composability. Such platforms can act as a receiver of data from the ubiquitous sensors, as a computer to analyze and interpret the data, as well as providing the user with easy-to-understand web-based visualization. Advanced algorithms are required to implement ML and AI approaches for big data analysis on computing clusters. A programming language suitable for working on big data (Python, R or others) could be used to write such algorithms or software. Therefore, a good knowledge of biology and IT is required to handle the big data from biomedical research. The most common among various platforms used for working with big data include Hadoop and Apache Spark.

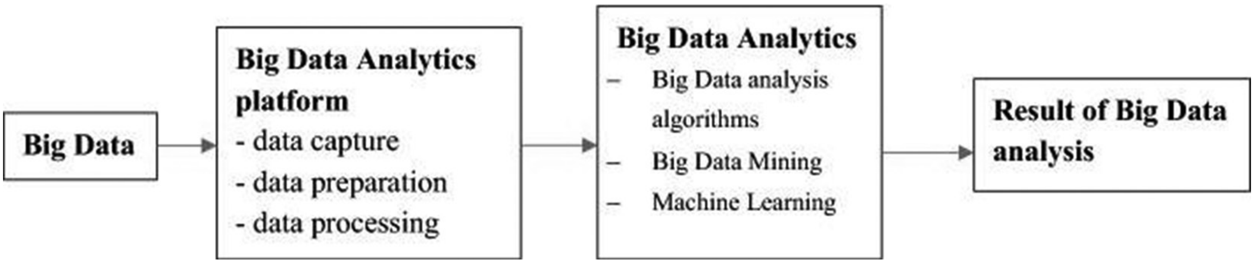

Fig.3: Big data analysis process

Hadoop and Apache Spark

Loading large amounts of bog data into the memory of even the most powerful of computing clusters is not an efficient way to work with big data. One of most popular open-source distributed application for this purpose is Hadoop. Hadoop implements MapReduce algorithm for processing and generating large datasets. MapReduce uses map and reduce primitives to map each logical record' in the input into a set of intermediate value pairs, and reduce operation combines all the values that shared the same key. It efficiently parallelizes the computation, handles failures, and schedules inter-machine communication across large-scale clusters of machines. Hadoop Distributed File System (HDFS) is the file system component that provides a scalable, efficient, and replica-based storage of data at various nodes that form a part of a cluster. It has other tools that enhance the storage and processing components therefore many large companies like Yahoo, Facebook, and others have rapidly adopted it. Many large projects, like the determination of a correlation between the air quality data and asthma admissions, drug development using genomic and proteomic data, and other such aspects of healthcare are implementing Hadoop. Apache Spark is another open-source alternative to Hadoop which is a unified engine for distributed data processing that includes higher-level libraries for supporting Spark SQL, streaming data, machine learning and graph processing.

ML for information mining, data analysis and estimates

In healthcare, patient data contains recorded signals for instance, electrocardiogram (ECG), images, and videos. Healthcare providers have barely managed to convert such healthcare data into EHRs. The unstructured, semi-structured and structured healthcare datasets have untapped wealth of information that can be harnessed using advanced AI programs to draw critical actionable insights in the context of patient care. In fact, AI has emerged as the method of choice for big data applications in medicine and smart system has quickly found its niche in decision making process for the diagnosis of diseases. Emerging ML and AI based strategies are helping to refine healthcare industry's information processing capabilities. For example, natural language processing (NLP) is a rapidly developing area of machine learning that can identify key syntactic structures in free text, help in speech recognition and extract the meaning behind a narrative.

Image analytics and IBM Watson

Some of the most widely used imaging techniques in healthcare include computed tomography (CT), magnetic resonance imaging (MRI), X-ray, molecular imaging, ultrasound, photo-acoustic imaging, functional MRI (fMRI), positron emission tomography (PET), electroencephalography (EEG), and mammograms who capture high-definition medical images of large sizes. Healthcare professionals like radiologists, doctors, nurses and others do an excellent job in analyzing medical data in the form of these files for targeted abnormalities. To help in such situations, image analytics is making an impact on healthcare by actively extracting disease biomarkers from biomedical images. This approach uses ML and pattern recognition techniques to draw insights from massive volumes of clinical image data to transform the diagnosis, treatment and monitoring of patients. It focuses on enhancing the diagnostic capability of medical imaging for clinical decision-making. A number of software tools have been developed based on functionalities such as generic, registration, segmentation, visualization, reconstruction, simulation and diffusion to perform medical image analysis in order to dig out the hidden information like Visualization Toolkit is a freely available software which allows powerful processing and analysis of 3D images from medical tests, while SPM can process and analyze 5 different types of brain images (MRI, PET, CT-Scan and EEG).

IBM Corporation is one of the biggest and experienced players in this sector to provide healthcare analytics services commercially. IBM's Watson Health is an AI platform to share and analyze health data among hospitals, providers and researchers. This platform utilizes ML and AI based algorithms extensively to extract the maximum information from minimal input. IBM Watson has been used to predict specific types of cancer based on the gene expression profiles obtained from various large data sets providing signs of multiple druggable targets. IBM Watson is also used in drug discovery programs by integrating curated literature and forming network maps to provide a detailed overview of the molecular landscape in a specific disease model. In order to analyze the diversified medical data, healthcare domain, are describe analytics in four categories: descriptive, diagnostic, predictive, and prescriptive analytics in IBM Watson.

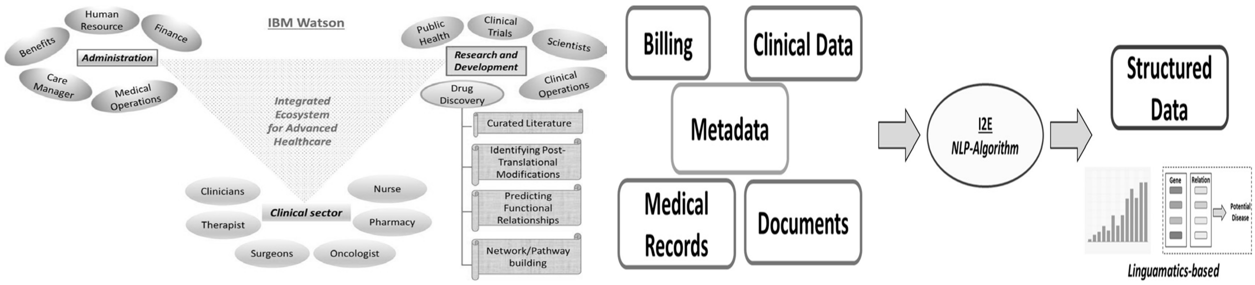

Fig. 4: Big data sources and tools

Challenges associated with healthcare big data

Methods for big data management and analysis are being continuously developed especially for real-time data streaming, capture, aggregation, analytics and visualization solutions that can help integrate a better utilization of EMRs with the healthcare. Ethical and legal challenges include the risk to compromise privacy, personal autonomy, as well as effects on public demand for transparency, trust and fairness while using Big Data.

Storage: Storing large volume of data is one of the primary challenges, but many organizations are comfortable with data storage on their own premises. It has several advantages like control over security, access, and up-time. It appears that with decreasing costs and increasing reliability, the cloud-based storage using IT infrastructure is a better option which most of the healthcare organizations have opted for. Organizations must choose cloud-partners that understand the importance of healthcare-specific compliance and security issues. Additionally, cloud storage offers lower up-front costs, nimble disaster recovery, and easier expansion and workable approach for providers with varying data access and storage needs.

Cleaning: The data needs to cleansed or scrubbed to ensure the accuracy, correctness, consistency, relevancy, and purity after acquisition. This cleaning process can be manual or automatized using logic rules to ensure high levels of accuracy and integrity. More sophisticated and precise tools use machine-learning techniques to reduce time and expenses and to stop foul data from derailing big data projects.

Image pre-processing: Medical images often suffer technical barriers that involve multiple types of noise and artifacts. Improper handling of medical images can also cause tampering of images for instance might lead to delineation of anatomical structures such as veins which is non-correlative with real case scenario. Reduction of noise, clearing artifacts, adjusting contrast of acquired images and image quality adjustment post mishandling are some of the measures that can be implemented to benefit the purpose.

Reduce Fraud, Enhance Security Meta-data: Some studies have shown that 93% of healthcare organizations have experienced a data breach. The reason is simple of personal data is extremely valuable and profitable on the black markets and any breach would have dramatic consequences. With that in mind, many organizations started to use analytics to help prevent security threats by identifying changes in network traffic, or any other behavior that reflects a cyber-attack. Of course, big data has inherent security issues and many think that using it will make organizations more vulnerable than they already are. But advances in security such as encryption technology, firewalls, anti-virus software, etc, answer that need for more security, and the benefits brought largely overtake the risks. There have been many security breaches, hackings, phishing attacks, and ransomware episodes that data security is a priority for healthcare organizations. After noticing an array of vulnerabilities, a list of technical safeguards was developed for the protected health information (PHI) help guide organizations with storing, transmission, authentication protocols, and controls over access, integrity, and auditing. Common security measures like using up-to-date antivirus software, firewalls, encrypting sensitive data, and multi-factor authentication can save a lot of trouble. To have a successful data governance plan, it would be mandatory to have complete, accurate, and up-to-date metadata regarding all the stored data. The metadata would be composed of information like time of creation, purpose and person responsible for the data, previous usage (by who, why, how, and when) for researchers and data analysts.

Visualization: A clean and engaging visualization of data with charts, box plot, heat maps, and histograms to illustrate contrasting figures and correct labeling of information to reduce potential confusion, can make it much easier for us to absorb information and use it appropriately like including bar charts, pie charts, and scatterplots with their own specific ways to convey the data.

Data sharing: Sharing data with other healthcare organizations would be essential. During such sharing, if the data is not interoperable then data movement between disparate organizations could be severely curtailed due to technical and organizational barriers. This may leave clinicians without key information for making decisions regarding follow ups and treatment strategies for patients. Solutions like Fast Healthcare Interoperability Resource (FHIR) and public APIs are making data interoperability and sharing easy and secure. The healthcare providers will need to overcome every challenge on this list and more to develop a big data exchange ecosystem that provides trustworthy, timely, and meaningful information by connecting all members of the care continuum. Time, commitment, funding, and communication would be required before these challenges are overcome. It would be easier for healthcare organizations to improve their protocols for dealing with patients and prevent readmission by determining these relationships well. Big data analytics can also help in optimizing staffing, forecasting operating room demands, streamlining patient care, and improving the pharmaceutical supply chain.

Information Quality: Information Quality (InfoQ), introduced in Kenett and Shmueli (2014), is a framework for planning, tracking, and assessing information derived from data and data analysis. It is formally defined as the utility, U, of applying a particular statistical analysis, f, to a particular data set, X, conditioned on a given goal. Given the four components: g, U, f and X, the formal definition is InfoQ = U(f(X|g)).

Besides these four components, InfoQ involves eight dimensions: Data Resolution, Data Structure, Data Integration, Temporal Relevance, Chronology of Data and Goal, Generalizability, Operationalization, and

Communication. Further developments of InfoQ and application examples are provided in Kenett and Shmueli (2016), Reis and Kenett (2018) and Kenett and Gotwalt (2021). The checklist of questions listed below can be used in evaluating the InfoQ dimensions. With reference to g, U, f and X we ask the following questions:

- *Data Structure*: Have you considered using data from different sources reflecting on the problem at hand?
- *Data Resolution*: Is the data granularity adequate for the intended goal? Has measurement uncertainty been evaluated and found appropriate?
- *Data Integration*: If relevant, how is the integration from different data sources done? Are there linkage issues that affect data privacy?
- *Temporal Relevance*: Does the time gap between data collection and analysis cause any concern?
- *Chronology of Data & Goal*: Are the analytic findings communicated to the right persons in a timely manner?
- *Generalizability*: Can you derive general conclusions based on the study, beyond what was explicitly studied, for example to other products or processes?
- *Operationalization*: Are the measured variables themselves of relevance to the study goal? Are stated action items derived from the study?
- *Communication*: Are findings properly communicated to the intended audience?
- The information quality framework is applied in the context of an information quality workflow (see Figure below). An information quality assessment can be performed at the design stage, during a project or in a retrospective evaluation of a completed study.

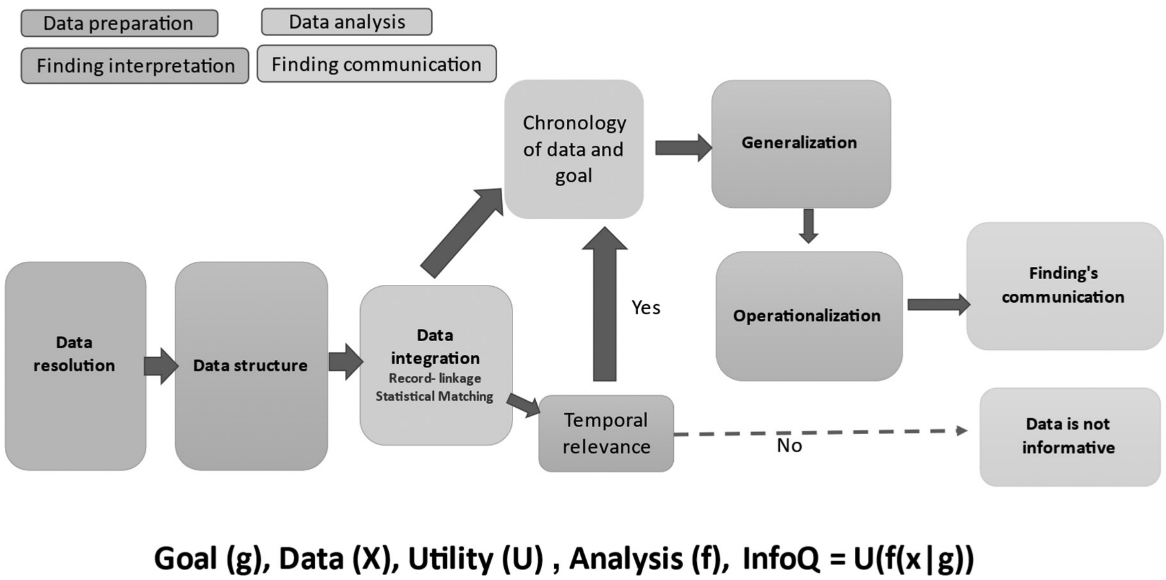

Fig.5: The information quality workflow

Quantum mechanics and big data analysis:

Big data sets can be staggering in size. Therefore, its analysis remains daunting even with the most powerful modern computers. For most of the analysis, the bottleneck lies in the computer's ability to access its memory and not in the processor. The capacity, bandwidth or latency requirements of memory hierarchy outweigh the computational requirements so much that supercomputers are increasingly used for big data analysis. The common digital computing uses binary digits to code for the data whereas quantum computation uses quantum bits or qubits. A qubit is a quantum version of the classical binary bits that can represent a zero, a one, or any linear combination of states of those two qubit states. Therefore, qubits allow computer bits to operate in three states compared to two states in the classical computation. Quantum computers use quantum mechanical phenomena like superposition and quantum entanglement to perform computations. Quantum algorithms can speed-up the big data analysis exponentially. Therefore, quantum approaches can drastically reduce the amount of computational power required to analyze big data.

Conclusions and future prospects

Nowadays, various biomedical and healthcare tools such as genomics, mobile biometric sensors, and smartphone apps generate a big amount of data. Big Data is beginning to revolutionize healthcare in the world as it offers paths and solutions to improve health of individual persons as well as to improve the performance and outcomes of healthcare systems. The analysis of such data can provide further insights in terms of procedural, technical, medical and other types of improvements in healthcare. After a review of these healthcare procedures, it appears that the full potential of patient-specific medical specialty or personalized medicine is under way. The collective big data analysis of EHRs, EMRs and other medical data is continuously helping build a better prognostic framework. Modern healthcare fraternity has realized the potential of big data and therefore, have implemented big data analytics in healthcare and clinical practices. Supercomputers to quantum computers are helping in extracting meaningful information from big data in dramatically reduced time periods. Clinical trials, analysis of pharmacy and insurance claims together, discovery of biomarkers is a part of a novel and creative way to analyze healthcare big data.

Big data analytics leverage the gap within structured and unstructured data sources. The shift to an integrated data environment is a well-known hurdle to overcome and big data relies on the idea of the more the information, the more insights one can gain from this information and can make predictions for future events. High volume of medical data collected across heterogeneous platforms has put a challenge to data scientists for careful integration and implementation. It is therefore suggested that revolution in healthcare is further needed to group together bioinformatics, health informatics and analytics to promote personalized and more effective treatments. Taken jointly, big data will facilitate healthcare by introducing prediction of, providing early warnings of disease conditions, and helping in the discovery of novel biomarkers and intelligent therapeutic intervention strategies for an improved quality of life.

References

Apple (2017): ResearchKit/ResearchKit: ResearchKit 1.5.3.

Dollas, A. (2014). Big data processing with FPGA supercomputers: opportunities and challenges. *In: 2014 IEEE computer society annual symposium on VLSI*.

Doyle-Lindrud, S. (2015): The evolution of the electronic health record. *Clinical Journal of Oncology Nursing, 19*(2),153–4.

Gillum RF. (2013): From papyrus to the electronic tablet: a brief history of the clinical medical record with lessons for the digital Age. *The American Journal of Medicine, 126*(10), 853–7.

Gubbi J, et al. (2013): Internet of Things (IoT): a vision, architectural elements, and future directions. *Future Generation Computer Systems, 29*(7), 1645–60.

Harrow, A. (2012). *Why now is the right time to study quantum computing*. XRDS.

Kenett, R.S. (2018), Quality of Information: a road map for applied data science. *The Voice of the System, 21*, 5.-21.

Kenett, R.S., & and Gotwalt, C. (2021) Maximizing Data Science Success with Information Quality (InfoQ) and JMP (2021-EU-45MP-750) JMP Discovery Summit Europe, https://community.jmp.com/t5/Discovery-Summit-Europe-2021/Maximizing-Data-Science-Success-with-Information-Quality-InfoQ/ta-p/349217

Kenett, R.S., & Maggino, F. (2021). Techniques for Analyzing and Presenting Official Statistics Indicators. *Statistical Journal of the IAOS, 37*(2), 1-12

Kenett, R.S., & Shmueli G. (2014) On Information Quality, *Journal of the Royal Statistical Society, 177*(1), pp. 3-38,

Kenett, R.S., & Shmueli, G. (2016) *Information Quality: The Potential of Data and Analytics to Generate Knowledge*. New York: John Wiley and Sons.

Kenett, R.S., & Zacks, S. (2021) Modern Industrial Statistics: With Applications in R, MINITAB, and JMP, 3rd Edition, New York: John Wiley and Sons.

Kenett, R.S., Zacks, S. and Gedeck, P (2022) *Modern Statistics: A Computer-Based Approach with Python*, Springer.

Laney, D. (2001). *3D Data Management: Controlling Data Volume, Velocity and Variety*. Stamford: META Group Research Note, 6.

Lloyd, S., Garnerone, S., & Zanardi, P. (2016): Quantum algorithms for topological and geometric analysis of data. *Nature Communications*, 7, Article number: 10138.

Mauro, A.D., Greco, M., & Grimaldi, M. (2016): A formal definition of big data based on its essential features.

Murphy, G., Hanken, M.A., & Waters, K. (1999). *Electronic health records: changing the vision. Philadelphia*: Saunders W B Co. p. 627.

Nielsen, M.A., & Chuang, IL. (2011): *Quantum computation and quantum information*. 10th anniversary ed. Cambridge: Cambridge University Press p. 708.

Raychev, N. (2015). Quantum computing models for algebraic applications. *International Journal of Scientific Engineering and Research, 6*(8),1281–8.

Reardon, S. (2017), Quantum microscope offers MRI for molecules. *Nature, 543*(162), https://www.nature.com/articles/nature.2017.21573

Reis, M., & Kenett, R.S. (2018) Assessing the Value of Information of Data-Centric Activities in the Chemical Processing Industry 4.0. *AIChE Journal, 64*(11), 42-48.

Reiser, S.J. (1991), The clinical record in medicine part 1: learning from cases. *Annals of Internal Medicine, 114*(10), 902-7.

Reisman, M. EHRs (2017). The challenge of making electronic data usable and interoperable. *Pharmacology & Therapeutics, 42*(9), 572–5.