Chapter 2
Big Data and Healthcare: Implications for Medical and Health Care in Low Resource Countries

Kgomotso H. Moahi
University of Botswana, Botswana

ABSTRACT

Big data and its application to healthcare has captured the world's imagination because of the ability of data analysts to combine huge disparate datasets and be able to produce trends, patterns and predictions. This ability lends itself to the quest to improve healthcare in terms of quality as well as cost. This chapter explores what big data is and how it can be applied to health care and medicine. To do this, the first sections address the question of what big data and data analytics are and what they encompass. An exploration of the potential benefits of big data is provided, with examples of applications, most of which are from the more developed nations of the United States and Europe. The chapter then considers what might be possible from implementing big data in low resource countries, with some examples of what already pertains. It looks at the challenges of implementing big data in health care in both developed and low resource countries.

INTRODUCTION

The data explosion that is being experienced in these times was long foretold by Vannevar Bush in his seminal paper of 1945. In that paper, he proposed an information storage and retrieval tool called the "Memex" which could function as a library for scientists and enable them to access information in a non-linear manner. To many people, his ideas must have seemed far-fetched, but time has proven him to be prescient. Digitization of many aspects of our lives has resulted in treasure troves of data waiting to be analyzed and used for business and other functions (Luna et al., 2014; Mayer-Schoenberger & Cukier, 2014; Manyika et al., 2011). Sectors such as banking and retail are realizing the value of big data to their businesses. Data from retail store loyalty cards is being analyzed for purchasing trends and patterns in order to determine marketing and delivery preference information. As various digital tools and applications are being used, "data exhaust" or "digital trail" is being generated, which lends

DOI: 10.4018/978-1-5225-2262-1.ch002

itself to a myriad of uses for innovation, productivity and competition (Manyika et al., 2011; Mayer-Schoenberger & Cukier, 2014), but also other less benevolent purposes. This data trail has immense potential in both private businesses as well as in the sphere of social services. The data is known as big data and its manipulation and use as big data analytics. Big data and its analysis have been touted as a means to generate information and knowledge in aid of decision making in various areas and domains. Healthcare is one of the domains typically covered in the literature (Asokan & Asokan, 2014; Austin & Kusumoto, 2016; Bellazi, 2014). Indeed public health in the USA has been cited as a beneficiary of big data analytics through Google Flu trends (Mayer-Schoenberger & Cukier, 2014).

Healthcare in general is information intensive and is characterized by immense data or record keeping (Raghupathi & Raghupathi, 2013). Information is critical for all aspects of healthcare such as patient care; public health, disease surveillance, decision support, and evidence based medicine. The use of computers in healthcare has a long history going back to the 1960s. Early applications of technology advanced as technology capabilities also developed. Much of the information used in health is digitized, for example electronic health records and clinical laboratory tests. Digitization of health data is not restricted only to developed economies (high income economies), but also to low and middle income economies where nascent e-health systems are being implemented. Furthermore, the move by many jurisdictions towards e-government means that even more information about their populations is coming on stream, and adding to the already burgeoning data trails and digital information.

Data produced in healthcare is characterized by high volumes, variety and velocity which are characteristics shared with big data (Belle et al., 2015). Big data is composed not only of structured data obtained from health information application, but also other data (unstructured) gleaned from various activities people engage in, in technology enabled services and applications – such as social media, location specific data enabled by sensors in mobile devices, searches on the Internet, and the content created on various applications. This type of data collection has been facilitated by rapid developments in technology, but has until recently not been harnessed and used effectively to improve health outcomes for a number of reasons. The reasons include the distributed and unrelated nature of the data, and the inability of traditional operations to manipulate the data (Belle at al., 2015). However, developments in cloud computing and data processing software such as Google's MapReduce and Apache Hadoop mean that more can be done with such data (Mayer-Schoenberger & Cukier, 2014).

Clinical data can now be combined with unstructured non-clinical data such as people's location and movement, lifestyles and behavior patterns to provide a composite image of the individual, and give healthcare providers the potential to influence lifestyle and behavior towards improved health outcomes. The use of big data analytics provides opportunity for data to be used to make predictions, to inform interventions, discern trends, and to ensure that the costs of health care do not escalate needlessly by addressing problems before they develop (Belle et al., 2015).

This chapter explores potential and actual uses of big data in healthcare and medicine. The literature is reviewed to define big data, big data analytics, and to describe some big data applications in healthcare. The review focuses on the benefits and the challenges of implementing big data in healthcare. The chapter also explores how healthcare and medicine in developing (low and middle income countries) might benefit from big data analytics and what the impediments to implementation are and how these may be addressed.