Abstract

In this paper, we analyze extant literature to assess the state of big data analytics relative to health care operations and report the preliminary findings of our research. Using multiple databases, we examine the literature to identify relevant resources addressing the health care areas of clinical operations, research and development, and public health. Big data analytics provides great potential for health care operations to increase patient safety and quality health care while reducing medical errors and costs. Additional benefits in health care organizations include improved quality of clinical decision-making, enhanced clinical trials and improved identification of public health hazard. Our research provides a foundation for future research by identifying current capabilities and implicating benefits to health care operations management professionals. Practitioners in health care operations gain from this study by recognizing means to increase patient safety and quality health care while reducing medical errors and costs.

Keywords (Required)

Big data analytics, operations management, clinical operations, public health.

Introduction

Big data analytics provides immense potential for health care operations managers to increase patient safety and quality health care while reducing medical errors and costs. Big data analytics also offer other benefits in health care organizations, including improved quality of clinical decision-making, enhanced clinical trials and improved identification of public health hazards. The purpose of this qualitative review of the literature is to explore big data analytics regarding health care operations, identifying current capabilities and implicating benefits to health care operations management professionals. With the increased use of health information technology as well as the digitization of historical records, health care organizations are creating substantial volumes of data. In this preliminary phase of study, we attempt to discern how organizations can use big data to improve health care operations over multiple disciplines, both now and in the future. Raghupathi (2014) notes the field of big data analytics as nascent and rapidly growing. Technological advances over the past decade have accelerated the growth of big data in health care, making this literature review unique as it provides implications for current and future practice in health care delivery and operations.

The remainder of the paper is structured as follows. First, using multiple databases, we analyze the literature to identify relevant resources addressing the health care areas of clinical operations, research and development, and public health. We then present our preliminary findings and discuss the implications of this study to health care operations practice and future research. We close with brief concluding remarks.
Method

We began our research with a search for relevant literature across the field of health care. We used the key words *big data health care* in the Google Scholar database and the consolidated online database (similar to Google Scholar) of a large mid-western university’s online library. Because the rate of technology change is rapid, we sought to use only very recent literature; hence, we constrained the search to include articles 2010 or newer. Additionally, we filtered the search to include only those publications listed as full text and that were from peer-reviewed journals. These criteria yielded 11,670 articles from the university database and 16,200 articles found in Google Scholar’s database. Then we used the following key phrases in Medline & ABI/Inform: *big data population health*, *big data health research*, and *big data clinical support*. We filtered the results to include articles and peer-reviewed journals since 2010, resulting in a yield of 4,740 articles.

Once we completed our data collection, we content analyzed the literature to extract relevant themes. Content analysis is a rigorous qualitative method used to discern meaning in various forms of communications media (Krippendorff 2004; Neuman 2006). To determine the state of big data analytics as it relates to the health care operations field, we used the process for problem-driven content analysis as advocated by Krippendorff (2004). At the time of this conference paper’s submission, only one of the researchers had coded the articles. As we continue our analysis, the others will also code the literature.

Findings

Although our study is ongoing, our analysis revealed interesting results. The past few decades have delivered the realization that health information technology (HIT) can greatly improve key operational concerns of quality and safety as well as those of reducing hospital errors – the cost of which is high (Gardner et al. 2015; Schneeweiss 2014; Tucker 2004). The increased use of HIT combined with the routine digitization of records and the incorporation of enterprise resource planning systems into health care facilities (Ward et al. 2014) is continuously creating massive volumes of data. The abundance of this electronically stored data, also known as big data, comes with the potential to create new knowledge about improved operations, the effectiveness of treatments, and predictability of outcomes (Schneeweiss 2014).

The rapid accumulation of data has created a relatively untapped resource with the power to change health care operations (Kamal et al. 2014). That being stated, there are examples of practitioners employing big data analytics in health care operations. One such example can be found in the Mayo Clinic’s use of a collection of data, its *Enterprise Data Trust*, to support the business intelligence (Chute et al. 2010). The Mayo Clinic integrated data from all aspects of its operation, including patient care, research, and administration, to aid in leadership decision-making (Chute et al. 2010).

Of importance for health care operations managers and providers is that with this vast amount of health data comes opportunity—opportunity to support health care functions across multiple settings. As health care professionals come to this realization they will begin to understand how big data not only greatly benefits them, but also their organizations and patients by increasing safety and quality of health care while concurrently reducing hospital errors and health care costs (Raghupathi and Raghupathi 2014). The awareness of the benefits of big data analytics is spreading and combined with assistance from big data information management platforms, these benefits are already becoming a reality (Raghupathi and Raghupathi 2014).

One of the essential components of making an accurate efficient decision is data (Abdelhak et al. 2014). Without data one cannot answer the necessary information questions referred to by Abdelhak and colleagues (2014) as to who, what, when, where, why, and how. Providing safe quality health care relies heavily on the accurate collection, analysis, interpretation, and reporting of data in a readable, usable format. To analyze the data, it likely will have to be modified to a useable format from what the source system provides (Ward et al. 2014). Providers need data to provide patients with evidence-based medical decisions; operations managers use error data to reduce hospital errors and costs of such errors; policy regulators use data trends to make population health decisions; and researchers use data for clinical trials (Abdelhak et al. 2014; Gardner et al. 2015).

Executives recognize the value of big data in markets across the world. Evidence of this is indicated by 85% of executives in Fortune 1000 companies stating they have big data plans and initiatives either
developing or currently in place (Barth and Bean 2012). Recognizing the value of big data has created a push for improved information technology infrastructures capable of collecting, analyzing, and interpreting vast amounts of data solely for the purpose of big data analytics and its associated benefits (Partington et al. 2014). Raghupathi and Raghupathi (2014) propose a big data information management architecture model that offers such potential benefits (Figure 1). Their model illustrates the cost and time reductions by streamlining the data-mining process, retrieving the minimum amount of data needed for the purpose at hand by employing algorithms and learners, and delivering that information in an immediate usable format through visual analytic tools. We believe this model suggests a foundational architecture on which to continue building. In the next section, we further examine current literature, specifically regarding the value of big data analytics across three health care functions: clinical operations, research and development, and public health.

**Clinical Operations**

As mentioned above, using data provides enhanced effectiveness in decision-making. Schneeweiss (2014) applies this concept to health care providers through two key learning applications of data, “the generation of new knowledge about the effectiveness of treatments and the prediction of outcomes” (p. 2161). Fortunately, his two applications of data are now becoming possible at a heightened state through big data analytic systems (Schneeweiss 2014). Additionally, Schneeweiss (2014) discusses the struggles that providers face when applying new medical knowledge to their patients because most evidence is based on studies and successes to patients that differ from their own. Imagine a system that retracts analyzed data from all over the country and filters that information to patient groups that match the specifications of a particular patient. Through a comparison of similar patients, a provider can feel assured they are studying actual care interventions instead of generic suggestions (Schneeweiss 2014). As more providers have success (or failures) in clinical operations and that data is collected, the effectiveness of a particular treatment method is more substantially validated, Schneeweiss (2014) continues. The original experts that developed the idea of clinical decision support systems wanted it to “think like an expert clinician when confronted with a patient” (Berner 2009). The sophistication of today’s big data analytics have made that a reality (Raghupathi and Raghupathi 2014).

As the volume of patient data in national and international databases increases, the amount of clinical support data that physicians have to combat rare diseases, similarly increases (Sessler 2014). These databases currently exist and are quickly accumulating large amounts of big data available to assist providers in clinical operations. Sessler (2014) lists several of these U.S. databases that aid in peri-
operative efforts such as the National Surgical Quality Program, Society of Thoracic Surgeons National Database, and the Multicenter Perioperative Outcomes Group Registry.

**Public Health**

Large national and international repositories of big data are available and open for epidemiological research (Genta and Sonnenberg 2014). Big data is essential in epidemiology to capture small variations in comparable populations. Genta & Sonnenberg (2014), have shown that new unsuspected associations are now identifiable through the use of big data analytics, (e.g., gastroenterology research). Digestive diseases are rare, with prevalence rates of less than 1% of the population (Genta and Sonnenberg 2014). Using traditional methods, it would be difficult to find a large enough cohort of patients with a specific diagnosis such as a rare digestive disease. However, using big data analytic tools, researchers can use filters and algorithms to identify a group of patients with the same demographics, risk factors, diagnosis, and the severity or stage of disease.

Ola and Sedig (2014) discuss the inclusion of data from local and/or national environmental and utility agencies into the all-data management information systems. This will allow a more complete picture of public health events and assist in the determination of contributing factors to health hazards. For example, the inclusion of water utility data will help understand the spread of diseases, such as cholera, through water distribution and consumption. Epidemiologist not only needs to be able to access data, but also must compare patient records, eliminate irrelevant data, apply environmental influences, and identify relationships among various factors. Big data analytics now provide this intense analysis identifying trends in a fraction of the cost and time as opposed to traditional methods (Sedig and Ola 2014).

**Implications**

Big data is more than a trendy buzzword. Big data enhances the operation of health care facilities, design and quality of clinical trials and research, and decisions on public health issues. Health care providers can use big data to accurately evaluate the tough outcomes and generate comparison groups that actually match patient demographics for effective case-control and retrospective cohort studies (Sessler 2014). Big data can also change the mindset of clinical researchers from a hypothesis-driven perspective to research based on analysis of the vast volume of existing data, similar to the evidence-based medicine approach that is gaining popularity among clinicians. The inclusion of data from other agencies outside of health care operations can assist epidemiologists with identifying unique health hazards. Big data allows physicians to go beyond current IT clinical decision support systems that merely offer providers reminders of follow ups, warning messages about drug interactions, and basic clinical guidelines in treatment of a disease (Berner 2009). Clinicians will be able to immediately extract information regarding comparable patients with similar diagnosis to determine accurate treatment methods. Current big data architecture can take this a step further by offering real-time input and use for providers around the world. As each success or failure of a treatment method is captured, the increased—or decreased—validity is automatically applied to each treatment method. We emphasize the importance of health care professionals using big data and promoting future big data systems to aid in important health decisions resulting in improved patient and public health.

**Conclusion**

Conclusions by policy makers, managers, and researchers that health data will greatly improve patient safety and increase quality of health care heralded the drive for digitizing and managing effectively health data. Advances in big data architecture—especially computer learning and visual analytic tools—have catalyzed the effective management and interpretation of vast amounts of data. These advances have made the theory of big data become a reality (Raghupathi and Raghupathi 2014), but health care operations professionals have still not derived the full advantage of this largely untapped resource.

Challenges exist in health care that limit the rate of expansion of big data analytics into practice. Frequently, health data are maintained in multiple, independent databases that—because of both proprietary software and privacy concerns—are not linked and have text fields that are not standardized...
(Genta and Sonnenberg 2014; Schneeweiss 2014). Health-related laws such as the Health Insurance Portability and Accountability Act of 1996 (HIPAA), with efforts to protect health information, constrain the level of adaptation of the big data stores (Weigel et al. 2009; Weigel et al. 2013). To overcome these challenges, future developers will have to evolve health care applications and systems to include more standardization; Health Level 7 International (HL7) is one such set of standards for communicating and sharing health information (Abdelhak et al. 2014).

Nonetheless, big data is here to stay. Its opportunity is seemingly endless and providers, researchers, epidemiologists, and health care operations professionals can use big data to ultimately improve patient care and safety. There is much the fields of operations and medicine have to teach one another (Abdulsalam et al. 2015; Boyer and Pronovost 2010). Health care providers and operations managers should support implementation of information technology systems that exploit big data and encourage their use as big data has proven to be a key player in achieving high quality, cost-efficient health care.

REFERENCES


