

# Issue Brief

*A Primer on Big Data Analytics (...and a Preliminary Q&A on the Basics of Big Data)*

June 2016

## Background

More and more clinical and administrative data are being digitized. New types of high volume data are being captured. Examples include genomics data and data streamed from monitors and sensors. Individuals are also tracking growing amounts of health data and sharing information about their lives via social networks. Furthermore, there is an ever-increasing range of clinical publications and reference sources.

Insights that could change health and care for individuals, groups, and even society as a whole are buried in these data. However, their complexity challenges common approaches to data storage, management, and analysis. The issue is not just the amount of data now available. It can also be overwhelming because of the velocity of data (frequency and speed of accumulation and sharing) and its variety. Highly diverse data – both structured and unstructured – are now the norm. Some experts also emphasize the difficulties posed by uncertain veracity of data, including the challenge of appropriately interpreting data in a ‘real world’ context.

Big data analytics (BDA) are designed to address these challenges. BDA involves new frameworks, technologies, processes, and capacities. It can be implemented in a variety of ways. For instance, both cloud-based and other types of BDA environments are possible.

## The Promise and Potential for Healthcare: Selected Examples

Globally, use of big data analytics in the health sector is still relatively new but growing. There are potential implications for all of the key aspects of healthcare organizations’ mandates.

### Clinical Care:

BDA has many potential applications in clinical care and supports the trend towards more person-centred care. For instance, Boston-based David Bates and colleagues identified six key opportunities to use predictive systems to stratify risks and predict outcomes: high

## Issue

Around the world, there is rapid growth in the volume, velocity, and variety of data. The result is ‘big data.’ It challenges traditional approaches to data storage, management, and analysis. Big data analytics, including predictive analytics, are mechanisms used to gain insight from patterns in these real-world data. They have garnered considerable attention in recent years and are beginning to be used in the health sector. [The promise and potential is strong](#). Early adopters have identified rich opportunities for improving population health, patient care, research, and education. At the same time, big data and big data analytics raise important questions and challenges for health organizations and policy-makers.

*This Issue brief was prepared  
for HealthCareCAN by:*

**Dr. Jennifer Zelmer**  
President, [Azimuth Health Group](#)

**HealthCareCAN has also  
appended a preliminary Q&A  
on the Basics of Big Data.**

cost patients, readmissions, triage, decompensation (when a patient's condition gets worse), adverse events, and treatment optimization particularly for patients with complex health problems.<sup>1</sup> Other healthcare applications include enabling personalized medicine, early warning systems based on data generated through remote monitoring, and analysis that supports management decisions (e.g. comparative effectiveness studies and identification of cost drivers).

### Population Health:

Big data is a natural fit to inform broad, based population health strategies. For instance, it can enable new approaches to disease and risk factor surveillance at individual, group, and community levels. By integrating data from different sources or perspectives, it is also possible to gain richer understanding of social determinants of health.

### Research:

BDA also expands the frontiers of health research. For instance, it can enable researchers to address new types of questions, to use new types of data, and to undertake larger-scale analysis than has historically been feasible. Where appropriate mechanisms are in place, it can also reduce the cost of research by enabling secondary use of data originally collected or generated for purposes other than research.

### Education:

New skills and capacities are needed if there is to be broad-based, effective use of BDA in the health sector. Highly skilled data scientists and analysts are essential. Patients/families, clinicians, managers, and policy-makers who seek to incorporate these approaches into their decision-making will also need to grow the ability to do so effectively. In addition, there are opportunities for educators to use BDA. For instance, it could help them to assess learner competencies.

## Key Challenges for Healthcare Organizations

The use of big data analytics is evolving rapidly, both in general and in the health sector specifically. Enablers and barriers to its effective use raise a number of challenges, including:

### Governance:

Updates to data, research, and other governance mechanisms – including frameworks, policy, regulations, and legislation – are required to fully leverage the possibilities of BDA and to ensure its appropriate use. This includes resolving issues related to data ownership and custodianship.

### Ethics:

Just because we can know, should we find out? A familiar question from fields such as genetic testing, it equally applies to a range of information that could come from BDA. These approaches also raise a number of other ethical questions.

For example, what is the obligation of someone who identifies a risk via BDA (e.g. a risk of suicidal ideation based on teens' social media posts) to act on that information?

### Privacy:

While existing privacy frameworks and legislation continue to have relevance, the breadth and potential of BDA stretch current privacy models and raise new issues. On-going review is needed. This may trigger a demand for new technological, policy, and organizational arrangements.

### Capacity/capability:

New infrastructure, partnerships, education and training, methodological advances, investments, and other developments may be needed to harness the opportunities that BDA offers.

### Business case:

Aligning value and understanding return on investment are likely to be central to scale and spread of advances in BDA.

### Clinical workflow:

Effective use of BDA to support clinical care may require new models of care and adjustment to current clinical workflows and practice patterns.

### Equity:

BDA is currently mostly used in larger, better resourced settings and its advantages may not be accessible to those served in other contexts.

### For More Information:

[\*Big Data Analytics in Health\* whitepaper](#), Canada Health Infoway

[\*Health Analytics for Informed Decision-Making: Health System Use Summit\*](#) (held in 2016 with presentations and panels by experts from Canada and other countries)

[\*Health Affairs July 2014 issue\*](#) (Special issue on the impact of big data on healthcare)

[\*Big Data Analytics in Healthcare: Promise and Potential\*](#) (includes a number of examples of applications in the health sector)

1. Bates D, et al. (2014). Big Data in Health Care: Using Analytics to Identify and Manage High-Risk and High-Cost Patients, *Health Affairs*, 33(7), 1123-31.

# A Preliminary 'Questions & Answers' on the Basics of Big Data

## What is Big Data?

Big Data refers to data that exceeds a sector's capacity for traditional means of collection, storage and analysis. As it is a relative term, there is typically no "threshold" for differentiating big data.<sup>1</sup>

## Why is it different from "data"?

Big data is differentiated by the volume, velocity, variability, and variety with which it is generated. Each person, record or encounter is a source of big data. The data may not be structured.<sup>2</sup>

## Is bigger always better?

Big data has important clinical, research and administrative potential. However, we have yet to exhaust the potential of regular datasets available to inform care and managerial decisions.<sup>3</sup>

## What are examples of the main sources of big data in healthcare?

Pharmaceutical R&D data (clinical trials etc); clinical data (EMR, images, monitors etc), activity and cost data (utilization etc), and patient and society (behaviours preferences, history, social media).<sup>4</sup>

## What accounts for the rise in Big Data?

In health, Big Data arises particularly from the impacts of the digitization of information, including information from electronic health records, genomics, mobile technology, social media, and other.<sup>5</sup>

## What infrastructure is required for Big Data?

Big data can require investment in technology, software, and platforms. Cloud computing can facilitate this by commodifying computing time and data storage through internet platforms.<sup>6</sup>

## What types of questions can we answer from Big Data?

"What happened" (descriptive analytics); "what will happen" (predictive analytics); "what did it happen" (diagnostic analytics) and "how can we make it happen" (prescriptive analytics).<sup>7,8</sup>

## What are sample policy issues that occur with Big Data?

Typical policy problems include who owns the data; will de-identification of the data be enough to protect individuals' privacy, do we have expertise to analyze and use the resulting findings.<sup>9,10,11</sup>

## What are examples of how Big Data enables clinical and research operations?

Comparative effectiveness; decision support; transparent medical data; remote patient monitoring; patient profiling; better modelling; clinical trials; personalized medicine; disease pattern analysis.<sup>12,13</sup>

## How can Big Data impact financial operations and new business models?

Big data can supply better cost and utilization information and facilitate economics and outcomes research, etc. It can allow for online communities; synthesized datasets, and surveillance.<sup>14</sup>

## References in the Q&A

1. Raghupathi, Wullianallur, and Viju Raghupathi. Big data analytics in healthcare: promise and potential. Health Information Science and Systems. February 7, 2014.
2. Ibid.
3. Ross, Jeanne W., Cynthia M. Beath, and Anne Quaadgras. You May Not Need Big Data After All. Harvard Business Review . December 2013.
4. Manyika, James, Michael Chui, Brad Brown, Jacques Bughin, Richard Dobbs, Charles Roxburgh, and Angela Hung Byers. Big data: The next frontier for innovation, competition, and productivity. McKinsey Global Institute. May 2011.
5. Janakiraman, Prakash, 2015. Big Data Cloud Database and Computing. Qubole. Available: <https://www.qubole.com/resources/?nabe=5695374637924352:1>
6. Ibid.
7. Hodder, Alyssa. How big data can influence health. Benefits Canada. May 29, 2015.
8. Iwashyna, Theodore J., and Vincent Liu. What's So Different about Big Data? A primer for Clinicians Trained to Think Epidemiologically. AnnalsATS. Volume 11. Number 7. September 2014.
9. Ross, et. al. You May Not Need Big Data After All. Harvard Business Review, . December 2013.
10. Manyika, et. al. 2011.
11. Raghupathi, et. al. February 7, 2014.
12. Manyika, et. al. May 2011.
13. Raghupathi, et. al. February 7, 2014.
14. Manyika, et. al. May 2011.