CANADIAN HEALTH INFORMATION MANAGEMENT ASSOCIATION

Professional Practice Brief
PPB – 0041.16

Big Data and Data Analytics

A professional practice brief consists of two major categories; both designed as professional development (PD) tools to advance health information professional practice and standards to support the delivery of quality healthcare. A PPB may relate to either category, or both. The two major categories are as follows:

☐ **Guidelines for Practice** ☑ **Professional Resource**

Table of Contents

October 2016

Big Data and Data Analytics

Introduction

Health data has grown in complexity over time. Health data comes from a variety of sources and formats – structured and unstructured, and requires sophisticated tools for complex analysis beyond common spreadsheets and relational databases. Given this complexity, health data can be described as “big data”. According to the SAS Institute, “big data is a term that describes the large volume of data – both structured and unstructured – that inundates a business on a day-to-day basis. But it’s not the amount of data that is important; it’s what organizations do with the data that is important. Big data can be analyzed for insights that lead to better decisions and strategic business moves”. The healthcare sector depends on big data in order to drive strategic outcomes. Big data exhibits a number of attributes. These are described based on the 4V’s - Volume, Velocity, Variety, and Veracity. Volume refers to the scale of data, Velocity - the analysis of streaming data, Variety - the various forms of data, and Veracity refers to the uncertainty of data. Advances in “data management, particularly virtualization and cloud computing, are facilitating the development of platforms for more effective capture, storage and manipulation of large volumes of data”. (See PPB 0042.16 on Cloud Computing). Big data requires a multidisciplinary approach using descriptive techniques and predictive models to derive valuable knowledge from data (analytics), and gaining insights that inform decision making (analysis).

Data and data analytics are essential in delivering quality healthcare to patients in many ways, some of which include clinical decision support, epidemic surveillance, and genomics (AHIMA, 2013). Data analytics can be very useful when it comes to clinical decision support, providing a means to make clinical decisions at the point of care.

Managing and analyzing big data requires the help of many individuals in a variety of roles. For example, it is necessary to have health information management professionals, big data analysts, researchers, etc., to successfully handle big data in the healthcare industry. Furthermore, there are skills required by each individual to work with big data successfully. Since big data is relatively new in the healthcare industry, there are several barriers in the use and analysis of big data. Lack of resources, privacy implications, and too much data are just some of the challenges that big data users must address relatively quickly.

Big data and data analytics are not limited to healthcare alone. In fact, the healthcare field lags behind other fields when it comes to analytics. One of the most robust fields that make extensive use of data analytics and big data is the retail industry. Retailers, both online and in store, have been tracking customer purchases for years. From this they have developed large databases of consumer purchases, requiring the use of data analytics to turn this data into information and knowledge (The Economist, 2010). This has allowed for personalized advertisements based on prior purchasing behavior and improved marketing strategies. Such marketing techniques have been highly successful for companies, demonstrating the value of big data analytics in today’s society.

Definitions of Big Data in Healthcare

When discussing big data in healthcare, the following two definitions come to mind:

“Large, complex, and relatively unstructured data sets, not suitable for analysis just using standard spreadsheet or relational database techniques” (Atkinson, 2014).

Big data gathers diverse pieces of information and already it has accounted for “improvements in personalized medicine, clinical risk intervention, waste and care variability reduction, automated external and internal reporting of patient data, and standardized patient registries” (Health Catalyst, 2016).

Big Data has been defined using the following criteria:

Volume: Volume refers to the sheer amount of data being collected. Millions of pieces of data are created, gathered, and stored everyday in the healthcare sector.

Velocity: Velocity refers to the rate at which data are generated. Not only are there large volumes of data being created, but these data are constantly being generated at an exceptional rate. “Data is being created in much shorter cycles, from hours to milliseconds” (Infoway, 2013).

Variety: Variety describes the different forms of data. Data generated in the healthcare field are stored in many different formats, due to the differing nature of the data itself. Images must be stored differently than patient records, which must be stored differently than ultrasound videos.

Veracity: Veracity refers to the uncertainty of data. This aspect of big data is critical in healthcare, as the decisions being made based on the data will have an impact on patients’ lives.

It is important to note, however, that big data refers not only to the data itself, but also the analysis techniques used to obtain information from such data.

Data Analytics Applications

There are numerous applications that have found use with respect to data analytics within healthcare, and the number of potential applications continues to grow as more data are collected from various sources. Data analytics applications include data from various facilities, including outputs from business intelligence and case-costing analysis. Data analytics includes secondary uses of data such as for research or population health monitoring. (See PPB 003R.16 on Health Data Access, Use, and Control for Secondary Uses).

Below are just a few of the applications of big data in healthcare. The number of applications of big data in healthcare is limited only by the data sources available and the creativity and expertise of the data analysts.

Clinical Decision Making

Electronic medical records (EMR) represent a potential source of information for making better clinical decisions including tools for: preventive screening, changes in lifestyle (diet and exercise), medications, etc. It is possible to identify trends that may indicate a problem that is shared by some portion of the population by looking at the data elements captured by the systems, such as morbidity or lifestyle factors. Analysis of these large health databases allows for the personalization of treatment plans based on the treatment and results of many similar patients. Examples of health recommendations stemming from healthcare data analytics are becoming more common. For instance, many health organizations have developed risk calculators that, when given information that an individual patient provides to them regarding their health and lifestyle, can calculate a risk score for the patient, and as well, provide recommendations to lower this score (Government of Canada, 2013). These risk calculators often rely on small datasets containing patient information to determine the impact of certain health and lifestyle factors on a patient's risk of developing disease. The use of the EMR, as a source of data, either locally or across a large network of multiple EMRs, could greatly expand the size of these datasets, allowing for better, more powerful predictors. EMR data represent an underutilized source of information that can be used in making better clinical decisions. Dr. Carolyn McGregor's Artemis project is an excellent example of how big data can play a role in clinical decision making. The Artemis system is able to take 1256 physiologic readings per second from premature infants and analyze these data to quickly predict adverse changes in the condition of the infant (IBM, n.d.). While these data are incomprehensible to a physician, the Artemis system is able to analyze the data to allow for earlier detection of life threatening conditions, such as sepsis, which in turn allows for faster, more effective treatment. This can often be the difference between life and death for these premature babies. From this example it is clear that big data is already helping physicians make better, more informed decisions and has the potential to expand even further into this area.

Genomics

Genomics examines the human genome, or an individual's entire DNA sequence, in an attempt to better understand disease (Mandal, 2014). Advancements in DNA technology have allowed for faster sequencing of DNA, creating extremely large databases of this genetic information (NIH, 2015). Big data and data analytics play a role in interpreting this data and what it means in regards to human disease. "Genome-based research is already enabling medical researchers to develop improved diagnostics, more effective therapeutic strategies, evidence-based approaches for demonstrating clinical efficacy, and better decision-making tools for patients and providers" (NIH, 2015). Clearly, one of the next steps in medicine is the ability to treat a patient based not only on the signs and symptoms of their disease (the 'phenotype' if you will), but also according to their genome. To do so requires extensive use of big data and data analysis techniques to gain knowledge from the vast genomic data resources. One such database is the Metabric database, housed by the Department of Molecular Oncology at the University of British Columbia. "METABRIC (Molecular Taxonomy of Breast Cancer International Consortium) is a Canada-UK project that aims to classify breast tumours into further subcategories, based on molecular signatures that will help determine the optimal course of treatment" (BC Cancer Agency, 2013). Work in this field is allowing researchers to determine the best treatment for a given tumour, which is not only improving patient outcomes, but also saving money on ineffective treatments.

Epidemic Surveillance

One application of big data is epidemic surveillance and the monitoring of infectious diseases. This has been seen in the use of data from either social media or search engines to predict epidemics and infectious disease outbreaks by monitoring trends in user inputs. Examples of these inputs would be keywords in search queries, and 'hashtag' key phrases in social media. A popular example of an epidemic surveillance program was Google's Flu Trends. Although they no longer publish estimates on flu and dengue fever, one can still access the historical data that were collected (Google, 2016). Google used the number of queries made by users to estimate the cases of flu and dengue fever based on geographic location. This was a successful project due to the fact that Google is such a widely used search engine. The data collected were very large in volume, and came in at a very high velocity. Google processes on average 40,000 search queries every second (Internet Live Stats, 2016). In their Google Flu Trends project, the estimation of occurrences were fairly good predictors. Google was able to map trends similar to those of the CDC, five to seven days before the agency released their own official predictions based on serotesting of cases (Dugas, AF. et al., 2013). Unfortunately, in 2012, Google Flu Trends predicted the peak of flu cases by over double the amount reported by authorities (Lazer, 2014). The project was ended due to this failure and no more data are being collected. Another example of using social media in Big Data Analysis is the Ailment Topic Aspect Model (ATAM) created to distinguish health topics from other topics on Twitter. The model was applied to data in 2009-2010 and was able to distinguish 15 different health topics. (Paul, 2011). Healthcare data could be used in a similar way to more accurately monitor population health by monitoring incoming patient symptoms and diagnoses.

Potential Users: Who deals with Big Data in healthcare?

In healthcare, different users will use big data and data analytics in different ways, based on what they wish to obtain from the data.

- **Healthcare providers** value big data analytics for timely access to patient, clinical, and other contextual data to improve their decision making and to provide efficient and safe care (Feldman, Martin, Skotnes, 2012).
- **Researchers** use big data to develop predictive models, statistical tools, and algorithms that can improve the quality of care and workflows for clinicians and other healthcare workers (Feldman, Martin, Skotnes, 2012).
- **Pharmaceutical companies** can optimize the usage of big data to gain a better understanding of the aetiologies of diseases, produce more effective pharmaceuticals, discover personalized medicine strategies in targeted drug therapeutics, and develop more successful clinical trials (Feldman, Martin, Skotnes, 2012).
- **Government officials** can use the information to reduce costs, enforce regulations, and maximize the social value of data by providing incentives to facilitate the adoption of EHR systems among providers (Feldman, Martin, Skotnes, 2012).
- **Health Information Management (HIM) professionals**, as the managers of personal health information who specialize in acquiring, analyzing, and protecting digital and paper-based health information, would benefit from big data to improve data quality and assist quality improvement programs within health facilities and health regions. Ultimately, this information is provided to planners, managers, and to support quality patient care (AHIMA, 2016). Current uses of data analytics for HIMs include:
  - Business analysis: answering questions regarding hospital operations using data (performance, quality assurance), and
  - Case-costing analysis: determining how much is spent on each patient during their stay.
- **Big data analysts**, such as business analysts, engineers, researchers, architects, statisticians, data governors, risk managers, or marketers, who design and execute analytics, are at the forefront of analyzing big data (Russom, 2011).

Skills required of HIMs interested in Big Data

Health Information Management professionals represent a unique and valuable source of personnel for data analytics; they have extensive knowledge of how the data are coded at a granular level and where the data are coming from. The vast majority of decision support units have one or more HIM professionals within their department. Other broad skills that are beneficial to HIMs wanting to explore data analysis include:

- Critical thinking skills: one must be able to develop ideas to be tested using data analytics and process improvement thinking
- Ability to make independent decisions
- Expertise in subject matter area allows analysts to ask the right questions
- Understanding of current events in healthcare
- Knowledge of privacy
- Knowledge of statistical analysis packages including SAS, SPSS, R, as well as more advanced tools such as Pentaho Business Analytics or Skytree Server
- Database management skills including SQL, Microsoft Access
- Strong communication skills: data analysis often requires input from people of different backgrounds. The ability to work with these people is essential.
- Skills in report writing to convey the knowledge gained from analysis

Skills required by users of big data demand a different approach and perspective. This includes strengthening the skills of clinical investigators and clinicians in data science, a new skill set that might become prominent in the near future (Krumholz, 2014). Chiefly, a change in mindset is required. One must be open to embracing machine learning, data mining, HADOOP, Extreme SQL, MapReduce, SaaS, and machine-based algorithms (Krumholz, 2014) (Russom, 2011). These computer-based tools assist in identifying patterns and relationships in data that might not be detected without their use. Competencies in informatics skills will also become necessary to understand how to best use big data (Krumholz, 2014). Furthermore, studies have shown that in the near future, the following skills will be essential: (Russom, 2011).

- **Advanced data visualization:** as data becomes more complex, methods of displaying such data must become more complex as well. Skills in advanced data visualization will allow analysts to effectively and dynamically present their findings using representations that go beyond simple histograms or bar charts. Interactive visual representations are an example of these advanced data visualizations, which allow the user to interact with the representation and gain insight into specific parts of the representation that are normally not accessible.
- **Real-time reporting:** real-time reporting allows users to determine the current status of a measured object instantly. This satisfies the need for data that is the most recent.
- **Text mining and natural language processing:** a large amount of data in an EMR is contained within the free-text notes taken by the healthcare professional (Nuance, 2012). Text mining and natural language processing give analysts the tools to extract information from these written passages without having to manually read through this text, something that is not feasible based on the huge volume of health records.
- **A working knowledge of advanced analytics tools** to allow for analysis of the complex and unstructured data that big data is composed of. A few tools, of many include: (Wayner, 2012).
    - **Pentaho Business Analytics:** this software is able to absorb information from the new data sources that big data is known for, such as highly unstructured data. Such software is able to access the data and present it as if it came from traditional SQL databases.
    - **Skytree Server:** Skytree Server takes less of a user-friendly visual approach, but with the right command, is able to execute sophisticated machine-learning commands. It rapidly clusters data, and then looks for outliers or data that are out of the ordinary that may be significant.
    - **Tableau Desktop and Server:** Tableau is a powerful visualization tool that allows users to visualize their data, then slice it up and visualize it in another way. Tableau allows you to work with Hadoop as you would with any other data source.

Infrastructure

Infrastructure, namely hardware and software, such as servers, databases, and applications, act as the core of the Big Data industry. They are used to handle the complex data that are processed, stored, and analyzed in the big data ecosystem (McNulty, 2014). In the past, relational databases were the major technology relied upon by enterprises to collect and process data. However, data have become much more complex and the volume, velocity, and variety of the data has significantly increased and can no longer be supported solely by a relational database (McNulty, 2014). Now, more advanced tools have been developed that have the ability to properly store, process, and analyze the enormous amounts of complex data that are dealt with today. Some of these infrastructural technologies include:

- **NoSQL**: A type of database that is able to process large volumes of multi-structured data, along with the discrete data stored among that multi-structured data (McNulty, 2014).
- **Hadoop:** A network of technologies specifically designed for the processing, storing, and analyzing of data. The core structure of Hadoop technologies is focused on separating and distributing data into parts and then analyzing those parts simultaneously, rather than trying to tackle single blocks of data (McNulty, 2014).
- **Massively Parallel Processing (MPP) Databases:** The basis of MPP databases is to segment data amongst multiple nodes, and then process these data segments in parallel, using SQL (McNulty, 2014).

In order for Big Data projects to be completed to their highest standards, enterprises need to make major investments in these software and infrastructural technologies, as well as into training existing personnel on big data technologies and how best to use them (Vanson Bourne, 2015).

Barriers to Implementation

There are several barriers to implementation that stand in the way of big data becoming widespread.

Too much data: Big data technologies in commerce and finance are different from technologies utilized in healthcare. “Despite significant technological progress, the big data revolution in healthcare is still in its infancy” (Jadhav, 2016). Data in healthcare are combined in such large quantity that it is difficult to figure out what data are relevant and what should be used. “Like analysts in other industries, healthcare professionals struggle to bring together masses of data and synthesize it into actionable information” (Jadhav, 2016). There is a lack of understanding as to what part of the data that are collected should be used and how valuable that information is. “When it comes to data analytics and big data in healthcare, many organizations struggle to understand the data, never mind analyzing it and reaping the benefits” (Lee, 2015).

Lack of resources: One of the largest impediments to the full implementation of big data in healthcare is a lack of resources, which includes both financial and human resources. Despite recent growth in these areas, the current financial resources allocated to decision support departments are not sufficient to realize the optimal potential of data analytics (Neumann, Vandervecht, 2016). A lack of skilled individuals also reduces the potential that data analysis is able to reach. Much of the training of employees must be done on the job, often requiring two years before an employee is up to a fully functioning level (Neumann, Vandervecht, 2016).

Healthcare regulations/Privacy: Healthcare regulations often limit what data are stored, as well as how and by whom these data can be used. “As big data analytics (BDA) progresses within the Canadian healthcare system, challenges, risks and concerns will arise about privacy, ethics and the legal implications of BDA. These risks and challenges represent significant barriers and potential drawbacks which may impact the success, pace and rate of adoption of BDA in Canada and the potential benefits it can bring” (Infoway, 2013). Classical privacy policies are commonly based upon three main principles: transparency, limiting use plus consent, and minimization (Stoddart, 2013). However, big data has forced this privacy model to be completely re-organized. For big data to be collected effectively, users cannot be asked to consent each time that the collected data are used for another purpose. On that note, the purpose for which the data is collected is not always known upon collection, so the user does not give direct consent for the use of their data in specific scenarios. Due to the differing research topics or analysis goals, the data collected are also normally not minimized, instead, the analysts will draw what information they need from the entirety of the stored data. Balancing innovative technologies with patient privacy is a problem that big data within healthcare faces and is a major barrier to widespread implementation. Privacy policy standards should always be taken into consideration when conducting big data projects or creating technologies that use or deliver big data analytics. However, adequate consideration needs to be given to the quality of the data so that the patients and health system benefit from the big data analysis.

A good example of a health-data-specific privacy policy standard is Privacy by Design (PbD), which was started in 1990 by the Ontario Information and Privacy Commissioner (IPC, 2016). It consists of seven fundamental principles that should be considered when using patient data. To highlight a couple, the first principle states that technologies should be proactive and preventative as opposed to reactive and remedial. Privacy issues should be anticipated and protected before-the-fact, not after (Cavoukian, 2011). The third principle states that privacy should be embedded into the design, and not added on as an extra feature (Cavoukian, 2011). For example, a technology that allows users to mine patient data should already have de-identified the personal information.

Future Directions

The healthcare industry will continue to generate data and depend on analytics not only to make sense of the data, but to drive important strategic outcomes. Through the process of "digitizing, combining and effectively using big data, healthcare organizations ranging from single-physician offices and multi-provider groups to large hospital networks and accountable care organizations stand to realize significant benefits". (Raghupathi, Raghupathi, 2014). To effectively take advantage of the possibilities of big data and analytics in healthcare, policies, standards and clear guidelines will need to be put in place. The healthcare industry is just beginning to see the value of big data that other industries have been developing for years. The benefits of big data and data analytics are numerous, and the use of healthcare data for analytic purposes is inevitable. As the people dealing with health data on a daily basis, health information managers are positioned to be at the forefront of this developing field, given the right tools and skills.

References

A.F. Dugas, M. Jalalpour, Y. Gel, S. Levin, F. Torcaso, T. Igusa, R.E. Rothman (2013). Influenza forecasting with Google flu trends. *PLoS ONE, 8* (2013), p. E56176

AHIMA (2016). Health Information 101. Retrieved February 29, 2016 from http://www.ahima.org/careers/healthinfo

Atkinson, G. (2014). Big data-what is it and what use is it? *The Journal of Ambulatory Care Management, 37*(3), 196-198. doi:10.1097/JAC.0000000000000039

BC Cancer Agency. (2013). *Metabric*. Retrieved from Department of Molecular Oncology, University of British Columbia: http://molonc.bccrc.ca/aparicio-lab/research/metabric/

Bipartisan Policy Center. (2013). *A Policy Forum on the Use of Big Data in Healthcare. Bipartisan Policy Center*. Washington: Bipartisan Policy Center.

Brimmer, Kelsey (2013). 5 ways hospitals can use data analytics. Healthcare IT News. Retrieved March 24, 2016, from http://www.healthcareitnews.com/news/5-ways-hospitals-can-use-data-analytics

Cavoukian, A. (2011). *Privacy by Design the 7 Foundational Principles*. Toronto, ON: Information and Privacy Commissioner of Ontario.

Feldman, B., Martin, EM, Skotnes, T (2012). Big Data in Healthcare Hype and Hope. *Dr. Bonnie 360*. Retrieved from http://www.west-info.eu/files/big-data-inhealthcare.pdf.

Google. (2016). *Google Flu Trends*. Retrieved from Flu Trends: About: https://www.google.org/flutrends/about/

Government of Canada. (2013). *Diabetes: Are you at risk?* Retrieved from Diabetes Questionnaire: http://healthycanadians.gc.ca/diseases-conditions-maladies-affections/disease-maladie/diabetes-diabete/canrisk/index-eng.php

Healthcare Analytics Adoption Model. (2016). *Health Catalyst*. Retrieved March 5, 2016, from http://health-information.advanceweb.com/Features/Articles/Big-Data-Barriers-in-Healthcare.aspx

IBM. (n.d.). *Smarter Healthcare in Canada: Redefining Value and Success.* Retrieved from Achieving Small Miracles from Big Data: https://www.ibm.com/smarterplanet/global/files/ca__en_us__healthcare__smarter_healthcare_data_baby.pdf

Infoway. (2013). *Big Data Analytics in Health.* Retrieved from Canada Health Infoway: https://www.infoway-inforoute.ca/en/component/edocman/resources/technical documents/emerging-technology/1246-big-data-analytics-in-health-white-paper-full report?Itemid=101

Internet Live Stats. (2016). *Google Search Statistics*. Retrieved from Trends and More: http://www.internetlivestats.com/google-search-statistics/

IPC - Office of the Information and Privacy Commissioner/Ontario | Introduction to PbD. (n.d.). Retrieved March 25, 2016, from https://www.ipc.on.ca/english/Privacy/Introduction-to-PbD/

Jadhav, S. (2016). Big Data Barriers in Healthcare. *Advance Healthcare Network*. Retrieved March 5, 2016, from http://healthinformation.advanceweb.com/Features/Articles/Big-Data-Barriers-in-Healthcare.aspx

Krumholz, H. M (2014). Big Data And New Knowledge In Medicine: The Thinking, Training, And Tools Needed For A Learning Health System. *Health Affairs 33.7.* 1163-170. Retrieved February 29, 2016 from http://content.healthaffairs.org/content/33/7/1163.full.

Lazer, D., Kennedy, R., King, G., & Vespignani, A. (2014). The parable of google flu: Traps in big data analysis.*Science, 343*(6176), 1203. doi:10.1126/science.1248506

Lee, K. (2015). Big data in healthcare: What's holding everyone back. *SearchHealthIT*. Retrieved March 7, 2016, from http://searchhealthit.techtarget.com/feature/Big-data-in-healthcare-Whats-holding-everyone-back

Mandal, A. (2014). *What is Genomics?* Retrieved from News Medical: http://www.news-medical.net/life-sciences/What-is-Genomics.aspx

McNulty, E. (2014, June 19). Understanding Big Data: Infrastructure. *Dataconomy*. Retrieved March 07, 2016, from http://dataconomy.com/understanding-big-data-infrastructure/

Neumann, Lila; Vandervecht, Agnes. (2016). Personal Communications.

NIH. (2015). *A Brief Guide to Genomics*. Retrieved from National Human Genome Research Institute: https://www.genome.gov/18016863

Nuance. (2012). *The Electronic Patient Narrative.* Retrieved from http://www.nuance.com/ucmprod/groups/healthcare/@web-enus/documents/collateral/nd_002650.pdf

Paul, M., Dredze, M. (2011). *A Model for Mining Public Health Topics from Twitter.* Technical Report. John Hopkins University.

Raghupathi, W., & Raghupathi, V. (2014). Big data analytics in healthcare: Promise and potential. *Health Information Science and Systems, 2*(3). Retrieved October 23, 2015, from http://www.biomedcentral.com/content/pdf/2047-2501-2-3.pdf

Russom, P. (2011). Big Data Analytics. *TDWI Research*. Retrieved from http://www.tableau.com/sites/default/files/whitepapers/tdwi_bpreport_q411_big_data_analytics_tableau.pdf

SAS Institute. Retrieved from http://www.sas.com/en_us/insights/big-data/what-is-big-data.html

Shah, P. (2016). The 4 “Vs” of Big Data Analytics in Healthcare. Retrieved from https://www.linkedin.com/pulse/4-vs-big-data-analytics-healthcare-parvez-shah

Stoddart, J. (2013). *The Age of Big Data: A new, challenging frontier for privacy and our society*. Retrieved from Office of the Privacy Commissioner of Canada: https://www.priv.gc.ca/media/sp-d/2013/sp-d_20131017_e.asp

The Economist. (2010). *Data, data everywhere*. Retrieved from Special report: Managing information: http://www.economist.com/node/15557443

The Four V's of Big Data. (2016). *IBM Big Data Hub*. Retrieved March 5, 2016 from http://health-information.advanceweb.com/Features/Articles/Big-Data-Barriers-in-Healthcare.aspx

Vanson Bourne (2015). The State of Big Data Infrastructure: Benchmarking global Big Data users to drive future performance. Retrieved March 07, 2016 from http://www.ca.com/content/dam/ca/us/files/industry-analyst-report/the-state-of-big-data-infrastructure.pdf

Wang, Y., Kung, L., Wang, W., & Cegielski, C. (2014). Developing a Big Data-Enabled Transformation Model in Healthcare: A Practice Based View. *IS in Healthcare. Thirty Fifth International Conference on Information Systems.* Retrieved February 29, 2016, from http://papers.ssrn.com/sol3/papers.cfm?abstract_id=2519257

Wayner, P. (2012). *7 top tools for taming big data*. Retrieved from InfoWorld: http://www.infoworld.com/article/2616959/big-data/7-top-tools-for-taming-big-data.html?page=2

Yaqub, F. (2013). Transparency in reporting clinical trials. *The Lancet Oncology, 14*(11), 1051-1053. doi:10.1016/S1470-2045(13)70469-8

Special Thanks to the Authors and Reviewers

Authors:

Nimi Bassi, BESc
Grace Bennett, BESc
Jason Black, BMSc
Navgeet Kaur Heran, MBBS
Alberto Lee, BHSc

Reviewers:

Andy Igonor, PhD, CISM, CRISC, CPHIMS
Michael Levy, BComp., PhD
Lila Neumann, BEng

Editor:

Paula Weisflock, BHA (HIM), MAEd, CHIM

For more information:

To learn more about this or any other Professional Practice Brief (PPB), including how to earn continuing professional education (CPE) credits from their use, please contact the Canadian Health Information Management Association (CHIMA) at:

Email: chima.education@echima.ca
Tel: 519-438-6700 Toll Free: 1-877-332-4462
Fax: 519-438-7001 Website: www.echima.ca

The publication of professional practice briefs (PPB) is a lengthy process due to the time it takes for authors to complete their research and write a draft of the brief. Subject matter expert reviewers are then called upon to review the PPB and provide feedback to the authors to consider incorporating into their work. Information within the PPB is deemed current at the time of writing; new information is always coming available and readers should take this into account. The content within this PPB reflects the opinions of the professionals who developed it. Scientific validation through research is not a part of the process. The links within this publication were current and valid as of the date of creation; however, webpage addresses and the information on them may change at any time.