COMMENTARY
Volume 8 Issue 1 2016
DOI: 10.5959/eimj.v8i1.417
ARTICLE INFO
Submitted: 21-02-2016
Accepted: 10-03-2016

Big Data: Applications in Healthcare and Medical Education

Fatimah Lateef

Department of Emergency Medicine, Singapore General Hospital

ABSTRACT

Big data is a hot topic these days, especially in academic circles. However, it is really a topic that goes beyond the four walls of academia. It has got practical applications in day to day practice of medicine and in medical education. What are these and are we ready to embark on the journey to continue to enhance the way we practice, into the future.

Keywords: *Big data, Analytics, Healthcare, Medical education, Electronic medical records*

CORRESPONDING AUTHOR Associate Professor Fatimah Lateef, Department of Emergency Medicine, Singapore General Hospital, Outram Road, 1 Hospital Drive, Singapore 169608 | Email: fatimah.abd.lateef@singhealth.com.sg

Introduction

The amount of new information created in the world is tremendous. It was estimated that in 1999, there was some 1.5 billion gigabyte of information, which doubled by 2003. Today this is increasing at an even more rapid rate and the number of "searches" using "big data" reached its peak in 2015 (1, 2).

So how big exactly is big data? It represents data of a very large size, typically to the extent that its manipulation and management presents significant logistical challenges. The datasets size is beyond the ability of typical database software tools to capture, store, manage and analyse. The data management challenges are often classified as the "4 V", that is volume, variety, veracity and velocity challenges. In the earlier years, there was no suitable software that could handle and manipulate such a large amount of data, but today, tools are available to analyse trends and identify patterns even in very large populations. Information can be readily connected to establish relationships and correlations (1–4).

So much data is generated daily, in a variety of industries, including healthcare. Whilst many other industries have been harnessing the value of large scale integration and analysis of big data, it is still considered relatively new in healthcare. Essentially, what can be measured, can be collected as data. However, it is also about how individuals, groups and organisations utilise and share data to effect necessary changes and impact population. At the same time, maintenance of confidentiality and data protection are major considerations (4–6).

Big Data and Healthcare

Big data collected in healthcare has many applications. Electronic medical records in institutions can generate a tremendous amount of data. Hospitals, laboratories, insurance companies and even the pharmaceutical giants are actively gathering data related to cost, diseases, medications, and warehousing them. These are valuable data and when analysed, can generate trends, observations, sub-group findings, precursors of diseases, genetic information

and many more. These can have impactful use in public healthcare decisions and planning, medications and their efficacy, amongst others. When securely linked and correlated, these heterogeneous data sets can have tremendous potential and applications in healthcare and health promotion. These information and data are also useful to assess healthcare cost, affordability of care and help plan enhancements in the quality of care (3–6).

Some do have the opinion that big data is "dirty data" and may not be robust enough for adequate utilisation and decision making (1, 3). The lack of standardisation is one such consideration. The various forms in which data is generated in healthcare may also pose challenges. This can range from the doctors' handwritten or electronic notes, to images, to insurance claims and even social media commentaries. The data collectors are also very heterogeneous and would include healthcare providers, administrators, nurses, students, statisticians, personnel from healthcare related companies and even other wellness institutions, as well as pharmaceutical companies. This adds to the variability. Therefore, the way it is used and its interpretation becomes critical. Integration will require collaborative leadership from both the public and private sectors. The context of its usage and application also becomes important. This also brings us to the issues of privacy, confidentiality, data protection and ethics. Much of the data captured electronically on our patients does not involve formal consent. Medical records are usually governed by rules that protect privacy in ways that traditional consumer marketing data is not. These are also indirectly bounded by the code of practice of doctors and researchers. Even if we argue that these medical data can be anonymised and have all identifiers removed, there is always the risks of breaches. As long as one person somewhere out there can perceive that some application is with reference to them, we are touching the ethical grey area, with the risk of privacy or confidentiality being compromised (1, 3–6).

Healthcare organisations thus have to come up with standards and frameworks to direct data governance and usage. It is not uncommon these days to have institutions employing data analytic specialists, Information and Technology Quality officers and other similar positions to help alignment. Insurance companies dealing with healthcare claims to use big data for their specific needs. They perform data and risk pooling to assess claims and decide on premiums for universal healthcare coverage for example (5, 6).

Big Data and Medical Education

The evolution seen in healthcare reflects the need to adapt to the changing role of technology and data. Teaching and inculcating these capabilities and skills thus becomes critical. Mindset change is also not just for our students and undergraduates, but also for practitioners, faculty and healthcare administrators. This way, the gap between how our learners are trained and how they practice and deliver healthcare can be narrowed. The ability to interpret and analyse these data is important for these new doctors and personnel. The teachers and faculty too will need to be able to negotiate this with their lifelong learning and continuing education as well as research works. With more objective use of data, it is the intention to have more personalised care and enhance the practice of precision medicine (6–8).

In medical education, teachers and faculty have to be empowered and trained on the use of big data. There should be dedicated time for this in order to get buy in and concurrence. Resources and open channels of communications must also be established. Students and practitioners will find data handling capabilities useful for their studies, research and projects. For example, with longer term and longitudinal studies which require follow up, how will the data be extracted and stored, before analysis. Being

IT savvy will also save time and enhance efficiency of performing these studies. Data such as these can be kept abreast of trends and new observations. The spectrum of healthcare personnel can also learn the utilisation of artificial neural networks and high performance computing when they are at a more advanced level (9, 10).

Today, every first and second year student at New York University Medical School is required to do a "healthcare by the numbers project". Students are given access to a database with some five million anonymous records. The database include age, race, ethnicity, diagnoses, procedures, bills paid for on their behalf and so on. Students are then to use tools provided to look at quality measures for diagnoses such as heart failure, hypertension and diabetes mellitus. They can categorise and create sub-groups as necessary too. The class it seems is appealing not just to data "junkies". The students are taught about looking at the practice of medicine from a different perspective and view (11). For example, the students have noted that big data is very useful with medical forecasts and setting up of a pharmaco-vigilance system for tracking adverse reactions and errors in practice.

In carrying out research and studies, institutions must also ensure the Ethics Committee or the Institution Review Board (IRB) must be educated on the use of big data and the requirements. There may also be a need to review some of the existing guidelines to incorporate relevant changes so as not to be vague, leave a grey zone or doubtful guesses in the conduct of clinical trials and research studies.

Big data collected can be usefully transformed into an auxiliary instrument to support further educational development and improvement. It can provide insight and overview into the whole medical curriculum and practice. The analysis can meaningfully identify knowledge, skills and attitude constructs through alignment of teaching methods and assessments. Gaps and lapses between the intended and target curriculum when studied from these data, can provide evidence to either keep or change certain practices or subjects. This curriculum mapping must be done from the viewpoints of the students, faculty and even patients where possible. This can also be a good exercise for institutions to review teaching and learning methods (12–14).

Big Data and Integrated Healthcare

As countries are all moving towards Integrated Healthcare (IH), big data has surfaced at the appropriate time. IH is about linking multiple levels of care management, coordinating services and encouraging professional collaborations. It focuses on networks and connections, and often between separate organisations on the continuum of healthcare delivery. Common in approach is the significant initiatives to share information on quality, costs and even outcomes of care delivery. This is also about connected health, which is about a range and spectrum of information and collaborative technologies. Connected health is a major enabler of IH. This will be a major undertaking for communities and nations, and must thus, have coordination, clinical practice guidelines, ethical standards, privacy and confidentiality agreements. This will also open up a whole new research platform and even public-private partnerships in some cases (14, 15).

Moving into connected health, the prerequisites would include (16):

1. Having a digital or IT infrastructure. This can ensure clinical efficacy in patient care.
2. Controlled exchange of information between personnel and institutions in a regulated fashion. Proper use and approval algorithms too must be available.

3. Data analytics to derive practical and useful information which can be utilised for clinical decision making, population health decisions, drug and treatment delivery decisions.

Big data has the potential to even enhance the efficacy of randomised controlled trials, which can have a powerful impact on connected health and integration of health systems. This can help us gain a better insight into what works and for whom or which groups of people and patients. This aligns with translating knowledge into actual practice. This will be evidence-based.

Handling these large quantities of data to generate such observations and results does not come easy, nor cheap. Investments in the appropriate systems and skilled manpower across the sector is crucial. These days, the use of predictive analytics, graphical and visual analytics, natural language processing, artificial intelligence networks and techniques for harvesting unstructured data are available for utilisation by trained staff. They can help health practitioners and administrators understand what we are dealing with and what we will need to deal with in the future. It is an exciting venture as more and more countries strive to become smart and connected nations (8, 9, 14).

Conclusion

If the data is big enough and the search exhaustive enough, the patterns can indeed be compelling enough. For the big data initiative to succeed, our healthcare systems must undergo some fundamental changes, not just in hardware and IT usage and acquisition, but also with the people in these institutions. Much of the potential for value from big data remains unclaimed. The venture and opportunities are exciting, in this age of "data liquidity". There are promising threads of knowledge to be discovered. It is entirely up to us and our next generation, to reap the rewards.

References

1. Gill Press. 12 big data definitions: what's yours? Forbes. [cited 2016 Feb 20]. Available from: http://www.forbes.com/sites/gilpress/2014/09/03/12-big-data-definitions-whats-yours/#305ec8b621a9

2. Google Trends. Available from http://www.google.com/trends/?hl=en-GB

3. Frieden J. The changing face of medical education: big data and self-paced learning. Med Page Today. [cited 2016 Feb 20]. Available from: http://www.medpagetoday.com/PublicHealthPolicy/MedicalEducation/52990

4. Buie S. Big data, better health. The Innovation of Care. 2014 Jan 29 [cited 2016 Feb 20]. Available from: http://www.theatlantic.com/sponsored/cvs-innovation-care/big-data-better-health/96/

5. Rosenbaum L. Transcultural chaos or enduring harm? The HER and the Disruption of Medicine. NEJM 2015;373(97):585–8.

6. Shah ND, Pathak J. Why healthcare may finally be ready for big data. Harvard Business Review. 2014 Dec 3 [cited 2016 Feb 20]. Available from https://hbr.org/2014/12/why-health-care-may-finally-be-ready-for-big-data.

7. Lee TH, Cosgrove T. Engaging doctors in the healthcare revolution. Harvard Business Review. 2014 June [cited 2016 Feb 20]. Available from https://hbr.org/2014/06/engaging-doctors-in-the-health-care-revolution.

8. Vaitsis C, Nilsson G, Zary N. Big data in medical informatics: improving education through visual analytics. Stud Health Technol Inform. 2014;205:1163–7.

9. Simpao AF, Ahumada LM, Galvez JA, Rehman MA. A review of analytics and clinical informatics in health care. J Med Syst. 2014;38(40):45. Epub 2014 Apr 3.

10. Triola M. Data and medical education: reimagining medical instruction. 2015 Sept 21 [cited 2016 Feb 20]. Available from http://evolllution.com/revenue-streams/market_opportunities/data-and-medical-education-reimaging-medical-instruction/

11. NYU School of Medicine. Healthcare by the numbers. 2014 [cited 2016 Feb 20]. Available from http://education.med.nyu.edu/ace

12. Ellaway RH, Pusic MV, Galbraith R, Cameron T. Developing the role of big data and analytics in health professional education. Medical Teacher. 2014;36(3):216–222.

13. Mennin S. Self organisation, integration and curriculum in the complex world of medical Education. Medical Education. 2010;44(1):20–30.

14. Baker RSJD, Yacef K. The state of educational data mining in 2009. A review and future visions. J Educ data Mining. 2009;1(1):3–16.

15. Olmos M, Corrin L. Academic analytics in a medical curriculum: enabling educational excellence. Australasia J of Educational Technology. 2013;28(1):1–15.

16. Accenture. Connected health: the drive to integrated healthcare delivery. 2012 [cited 2016 Feb 20]. Available from http://www.himss.eu/sites/default/files/Accenture-Connected-Health-Global-Report-Final-Web.pdf