HOSTED BY
ELSEVIER

Contents lists available at ScienceDirect

Journal of Acute Disease

journal homepage: www.jadweb.org

Review article http://dx.doi.org/10.1016/j.joad.2015.04.003

Big data as a new approach in emergency medicine research

Ho Ting Wong[1,2,3,4*], Qian Yin[2,3], Ying Qi Guo[5], Kristen Murray[4], Dong Hau Zhou[6], Diana Slade[4]

[1]*CUHK Jockey Club Institute of Ageing, The Chinese University of Hong Kong, Hong Kong Special Administrative Region*

[2]*State Key Laboratory of Resources and Environmental Information System, Institute of Geographic Science and Natural Resource Research, Chinese Academy of Sciences, Beijing, China*

[3]*Key Laboratory of Surveillance and Early Warning on Infectious Disease, Chinese Center for Disease Control and Prevention, Beijing, China*

[4]*International Centre for Communication in Healthcare, The Hong Kong Polytechnic University, University of Technology, Sydney, Australia*

[5]*Department of Social Work and Social Administration, The University of Hong Kong, Hong Kong Special Administrative Region*

[6]*School of Physical Education and Equestrian, Wuhan Business University, Wuhan, China*

ARTICLE INFO

Article history:
Received 13 Mar 2015
Received in revised form 17 Mar 2015
Accepted 10 Apr 2015
Available online 10 Jul 2015

Keywords:
Big data
Emergency
Medical research

ABSTRACT

Big data is a hot topic in the academic sector, and healthcare researchers are definitely not an exception. This article aims to provide a showcase in emergency medicine research to demonstrate the advantages of conducting such research using big data. Big data is a new and cost-effective research approach, and emergency medicine researchers could benefit from using this approach and by doing so producing high-quality research at a faster pace.

1. Introduction

Big data analytics is becoming increasingly popular. According to Google Trends, the number of searches using the keyword “big data” started to increase dramatically in 2011 and reached its peak this year[1]. Although the term “big data” sounds as if it is related to the area of data science only, it actually plays an important role in healthcare research, including emergency medicine.

Traditionally, researchers have adopted the 3Vs criteria to define big data: volume (*i.e.*, amount of data), velocity (*i.e.*, speed of data in and out), and variety (*i.e.*, range of data types and sources)[2]. However, from the view of socio-medical science, this definition may not be practical enough because the 3Vs criteria are dynamic and time dependent. For example, ten years ago, the volume of a DVD could be considered large, but now such a volume is common. Moreover, the size of the data is largely dependent on the data-compression technique adopted. Similarly, a 56K modem used to be considered fast, but broadband Internet speeds over 100M are ubiquitous in today's society. Another definition of big data “N = all” is much more practical[3]. According to this definition, researchers use big data if their data covers the whole population rather than only a sample of the population. However, outdated computing infrastructures are not robust enough to handle big data, making the big data approach not feasible. Moreover, clinical practitioners may not be comfortable with this approach because the majority of their research involves direct contact with patients. It is difficult to imagine a research project involving all the patients who fit the selection criteria. To save resources, clinical practitioners have to take a sample from the population and use inferential statistics to estimate the population's characteristics.

Compared with clinical practitioners, data scientists are more familiar with research that is focused on data. Particularly, analyzing very large numbers of hospital records of the whole

*Corresponding author: Ho Ting Wong, International Centre for Communication in Healthcare, the Hong Kong Polytechnic University & University of Technology, Sydney, Australia.
E-mail: frankwong@connect.hku.hk
Peer review under responsibility of Hainan Medical College.

population is one of the big data approaches that are becoming more and more popular. Population is the key for this to be a big data approach; otherwise, this will be considered a secondary data approach, which places its main emphasis on reusing existing data sets. The big data approach is cost effective because hospital records already exist; thus, the data-collection process can be avoided. Moreover, the quality of hospital records is expected to be more accurate than data obtained through questionnaire surveys.

2. Showcase

A review of the 204 papers published in the first eleven issues of the Journal of Acute Disease from 2012 to 2015 produced only one paper that resembles the big data approach[4]. In this connection, a showcase is particularly useful for encouraging clinical practitioners to conduct research using big data. Wong and Lai used big data containing over 6 million emergency department records to conduct research related to ambulance demand and weather[5–7]. They obtained the data set from the Hong Kong Hospital Authority, which contains patients' age, gender, triage level, and other information.

In their first published paper, Wong and Lai extracted ambulance-use records from the data set and aggregated them into time-series data by different socio-demographic and health factors[5]. The time series were then regressed on different meteorological time-series data, such as daily temperature and relative humidity, obtained from the Hong Kong Observatory (HKO) website[8]. Finally, high-risk groups—such as women, low-income groups, and the elderly—were identified to be more sensitive to extreme weather conditions.

In their second published paper, Wong and Lai used the same big data to develop a short-term daily ambulance-demand forecast system[6]. In addition to the daily ambulance-demand data series, the HKO seven-day weather forecast report was used to predict the next seven days' daily ambulance demand through the Autoregressive Integrated Moving Average model, available at IBM SPSS Forecasting[8,9]. Wong and Lai successfully demonstrated that the HKO seven-day weather forecast report is useful for improving the short-term daily ambulance-demand forecast. The results suggested that the ambulance managing authority could consider including such a forecast system in their Fourth-Generation Mobilizing System.

In their third publication, Wong and Lai used the ambulance-demand big data to develop a long-term projection of ambulance demand for the year 2036[7]. They used the big data and the population projection compiled by the Planning Department of Hong Kong to make the projection through combining regression models[10]. Their results are a warning signal for the government: without considering the rapidly aging population, the long-term ambulance-demand projection will largely be underestimated.

3. Conclusion

The three aforementioned studies demonstrate the potential of big data in emergency medicine research. Although only these three studies are shown, more can be conducted on the same set of big data to generate more research reports. This differs from traditional clinical studies, whose data sets normally can be used for only one study. Big data, by contrast, provides flexibility in analyzing a data set from different perspectives. Moreover, because governmental organizations already maintain a significant amount of big data, research using big data can be conducted at a very low cost. Because of the open government initiative by President Obama[11], it is now more popular for governments to release their big data for public use. In addition, because the data-collection process (*e.g.*, interview surveys) can be avoided, researchers who use governmental big data may complete their projects in a shorter time. Although most of the emergency medicine studies have to be conducted with traditional approaches, promoting big data as a new approach among clinical practitioners is worthwhile, particularly in the time short, and highly stressed contexts of emergency medicine.

Conflict of interest statement

The authors report no conflict of interest.

References

[1] Google Trends. [Online] Available from: http://www.google.com/trends/?hl=en-GB [Accessed on 19th February, 2015]
[2] Gartner. Gartner says solving 'big data' challenge involves more than just managing volumes of data. USA: Gartner. [Online] Available from: http://www.gartner.com/newsroom/id/1731916 [Accessed on 19th February, 2015]
[3] Mayer-Schönberger V, Cukier K. *Big data: a revolution that will transform how we live, work and think*. London: Eamon Dolan/Mariner Books; 2014.
[4] Haner A, Örninge P, Khorram-Manesh A. The role of physician-staffed ambulances: the outcome of a pilot study. *J Acute Dis* 2015; **4**: 63-7.
[5] Wong HT, Lai PC. Weather inference and daily demand for emergency ambulance services. *Emerg Med J* 2012; **29**: 60-4.
[6] Wong HT, Lai PC. Weather factors in the short-term forecasting of daily ambulance calls. *Int J Biometeorol* 2014; **58**: 669-78.
[7] Lai PC, Wong HT. Weather and age-gender effects on the projection of future emergency ambulance demand in Hong Kong. *Asia Pac J Public Health* 2015; **27**: NP2542-54.
[8] Hong Kong Observatory Website. [Online] Available from: http://www.hko.gov.hk [Accessed on 19th February, 2015]
[9] IBM. SPSS forecasting. United States: IBM. [Online] Available from: http://www-03.ibm.com/software/products/en/spss-forecasting [Accessed on 19th February, 2015]
[10] Planning Department Website. [Online] Available from: http://www.pland.gov.hk [Accessed on 19th February, 2015]
[11] The White House. Open Government Initiative. Washington: The White House. [Online] Available from: https://www.whitehouse.gov/open [Accessed on 19th February, 2015]