

## Big data and pain

Department of Anesthesiology and Pain Medicine, Asan Medical Center,  
University of Ulsan College of Medicine, Seoul, Korea

Jeong Gill Leem

“Big data” refers to the huge amount of digital information created that traditional databases are not equipped to store and analyze [1]. Big data includes medical, environmental, financial, geographic, and social media information and will continue to grow, added to by sources that are currently unimaginable.

The three Vs, volume, variety, and velocity, are normally cited when mentioning big data, and some authors add two more Vs: veracity and value [2,3]. Volume refers to the size of the data for processing and analysis. Velocity refers to the rate of the data growth and usage. Variety means the different types and formats of the data used for processing and analysis. Veracity concerns the accuracy of results and analysis of the data. Value is the added value and contribution offered by data processing and analysis.

In health and medical treatment, one of the main purposes of big data focuses on healthcare. Big data may stand to improve public health by providing insights into the causes and outcomes of disease, better drug targets for precision medicine, and enhanced disease prediction and prevention. Moreover, individuals will increasingly use this information to promote their own health and wellness, and to improve their understanding of health behaviors (smoking, drinking, etc.) and public or private healthcare [4].

Information added to big data is largely obtained from an official, convenient sample of people, as well as by informal means such as blogs, searches, and social media sites available on the internet. When searching on different databases such as Scopus, PubMed, Science Direct, Web of Science, and Cochrane Library with terms such as “pain” and “big data”, just a few related papers will be found. A possible explanation of this paucity of research results is that research in the field of pain using big data is still in its infancy. Another reason is that many researchers do not put “big data” as a key word or index despite using public data in their study.

As a kind of formal-official form of big data, public data refers to all databases, electronic files etc., that government and public sector organizations have electronically created or acquired and are managing. Actually, there are many papers written on the basis of the results analyzing the data from the primary care database, national hospital discharge records [5], the National (Health) Insurance database [6,7], Medicare and Medicaid Services, and the Physician/Supplier Procedure Summary Master data [8].

In Korea, recently, the Health Insurance Review & Assessment Service (HIRA) began to open the public data, a diverse and extensive collection of electronic medical records and resources, to researchers engaged in the

Received September 6, 2016. Accepted September 6, 2016.

Correspondence to: Jeong Gill Leem

Department of Anesthesiology and Pain Medicine, Asan Medical Center, University of Ulsan College of Medicine, 88 Olympic-ro 43-gil, Songpa-gu, Seoul 05505, Korea

Tel: +82-2-3010-3861, Fax: +82-2-470-1363, E-mail: [jgleem@amc.seoul.kr](mailto:jgleem@amc.seoul.kr)

© This is an open-access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (<http://creativecommons.org/licenses/by-nc/4.0/>), which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.

Copyright © The Korean Pain Society, 2016

Industry–University Institute Collaboration. The enormous open databases will enable various kind of large observational studies that examine the epidemics of disease, changing modalities of treatment, cost/utility–risk/benefit analysis, and rare but serious complications associated with invasive interventions for managing pain. Compared with randomized controlled trials, these studies have the advantage of access to quick, cheap, and easily obtainable information on a large population of patients in everyday clinical practice.

However, we should not overlook an important weak point of public data. Most of the information (data) in the public data was actually recorded for billing or documentation purposes such as insurance claims. It is just ‘found data’ which is not collected primarily for research. This renders the results of these studies susceptible to issues and biases not faced when dealing with traditional randomized controlled trials.

In order to overcome this weak point of found data, a consensus standardization of the data is required. For example, the incidence rate of postherpetic neuralgia is varied according to the definition of the duration of persisting pain (30, 90 or 180 days). The policy of privacy information protection and coverage range of the national health insurance service are the other obstacles in analyzing public data. Whenever analyzing big data, we should keep in mind that spurious correlations may lead to “big error”.

In conclusion, with advances in the technologies of collection, storage, transference, and state-of-the-art analytical methods, big data will greatly improve our

knowledge about disease management from diagnosis to treatment and prevention. As a pain specialist, I hope upcoming research will be focused on pain practice for developing decision making, practice guidelines, and, if possible, evidence based medicine.

## REFERENCES

1. de la Torre Díez I, Cosgaya HM, Garcia–Zapirain B, López–Coronado M. Big data in health: a literature review from the year 2005. *J Med Syst* 2016; 40: 209.
2. Chang V, Wills G. A model to compare cloud and non–cloud storage of big data. *Future Gener Comput Syst* 2013; 57: 56–76.
3. Huang T, Lan L, Fang X, An P, Min J, Wang F. Promises and challenges of big data computing in health sciences. *Big Data Res* 2015; 2: 2–11.
4. Khoury MJ, Ioannidis JP. Medicine. Big data meets public health. *Science* 2014; 346: 1054–5.
5. Gialloreti LE, Merito M, Pezzotti P, Naldi L, Gatti A, Beillat M, et al. Epidemiology and economic burden of herpes zoster and post–herpetic neuralgia in Italy: a retrospective, population–based study. *BMC Infect Dis* 2010; 10: 230.
6. Kim CH, Chung CK, Park CS, Choi B, Kim MJ, Park BJ. Reoperation rate after surgery for lumbar herniated intervertebral disc disease: nationwide cohort study. *Spine (Phila Pa 1976)* 2013; 38: 581–90.
7. Pannell WC, Savin DD, Scott TP, Wang JC, Daubs MD. Trends in the surgical treatment of lumbar spine disease in the United States. *Spine J* 2015; 15: 1719–27.
8. Manchikanti L, Pampati V, Falco FJ, Hirsch JA. Assessment of the growth of epidural injections in the medicare population from 2000 to 2011. *Pain Physician* 2013; 16: E349–64.