JACC: CARDIOVASCULAR IMAGING
VOL. 6, NO. 11, 2013

ISSN 1936-878X/$36.00
PUBLISHED BY ELSEVIER INC.
http://dx.doi.org/10.1016/j.jcmg.2013.09.007

# Are We Up to Speed?

## From Big Data to Rich Insights in CV Imaging for a Hyperconnected World

Jagat Narula, MD, PhD

The first software implementation of the heat map display was developed in 1991 (1) and has been used since then as a mainstream tool in systems biology for communicating interactions among genes, proteins, and metabolites that define life processes. In general, heat maps are considered a flexible visualization tool for data clusters and to explore patterns using rectangular, color-coded arrays that display the conceptual and statistical linkages among variables plotted along the x and y axes. This issue of *iJACC* presents a viewpoint in which heat maps are suggested as a step to explore the structure of big databases recorded during the assessment of cardiac function (2). The examples of data interactions that distinguish the pattern of left ventricular deformation seen in patients with constrictive pericarditis and restrictive cardiomyopathy illustrate the potential value of heat map visualization, symbolizing the burgeoning interest in big data in health care (3,4). Cardiac imaging and radiology are generating multiple terabytes of images every year, and the picture archiving and communication system market is projected to double in the next 5 years. Although several viewpoints have been recently published regarding the emerging role of big data in health care (3,4), it is worth questioning what healthcare big data one may possibly find concealed in the images and how this new knowledge could be extracted despite the existing challenges in healthcare information technology.

The term *big data* has evolved as a catchword for smart, more insightful data analyses of both structured and unstructured data. For cardiac imaging, big data analytics can refer to the information accrued from echocardiography, angiography, and magnetic resonance imaging or computed tomography. One could further contemplate the specific interactions of imaging data with electronic medical records, home monitoring devices, genomic data, insurance claims, and drug information for improving patient outcomes. Current economic concerns, perhaps more than any other factor, are driving more scrutiny and demands for big data applications. President Obama approved the $200 million Big Data Research and Development Initiative (5), and both the National Institutes of Health and the National Science Foundation have identified big data as a program focus for soliciting proposals that emphasize collection, deidentification, validation, archiving, and dissemination of large volumes of imaging and associated genetic, pathological, and clinical data generated from clinical research projects and clinical trials with the aim of enabling meaningful organization of the data. Furthermore, the United States federal government and other public agencies have recently made available their massive stores of healthcare knowledge, including data from clinical trials and information on patients covered under public insurance programs. The potential value of big data in health care and global economic trends in the utilization of mobile applications and telemedicine has also motived global organizations. For example, the United Nations has prioritized a workforce called Global Pulse that is currently working to connect the dots between data mining and humanitarianism, an effort that may enable digitally mapping the development of a global healthcare ecosystem. The big data revolution, indeed, has been rightly referred to as our "planet developing a digital nervous system"; several real-time sensors and monitoring devices are in the commercial pipeline and are expected to move diagnostics into an individual's own personal space, creating a hyperconnected world that liberates health-care monitoring and decisions from the walls of the clinic or hospital into individuals' communities and homes.

Although the prospects of big data in health care are exciting, its use in cardiovascular imaging poses unique challenges, both for individual providers and at the institutional level. At the provider level, some physicians are likely to feel intimidated by the decision-making prospects of big data science, arguing that intelligent machines and computational algorithms may displace the traditional art of medicine and critical thinking. Perhaps the appropriate attitude would be to form a human-computer synergy that will surpass the capabilities of either alone, just as we use Google Maps to

From the Icahn School of Medicine at Mount Sinai, New York, New York.

navigate our travel decisions or use the Global Positioning System for driving. There are other personal issues, such as the ownership of the data and the privacy of information that could lead to the identification of patients. This problem will require the development of robust Health Insurance Portability and Accountability Act–compliant interfaces and mechanisms for anonymizing the data analytics deluge.

Several challenges are likely to surface at the institutional level as well. First, institutions will be compelled to adopt more open-source platforms that allow the use of a large grid of inexpensive servers along with cloud-based services that support on-demand reporting and analytics. However, this change may be slow to adopt, paralleling the historical trends in the slow adoption of new health-care technologies in the United States (6). For example, when 88% of general practitioners in the Netherlands started using electronic medical records, only 17% of their American counterparts were seen to adopt these changes (7). These cultural disparities have contributed to the quality deficiencies in United States health care and would need to be addressed. While other industries, such as the financial, retail, and computer industries, have long-established uniform standards for replacing or updating their old technologies, health care usually lags behind. Often in institutions, one encounters myriad old existing homegrown systems and newer vendor systems. This, in turn, makes it very challenging to invest in a more advanced technology, not understanding the complex needs for networking with the older technologies. This lack of interoperability between different data network is rather rampant in health care and imaging and contributes to clinicians' resistance to adopt new systems. The growth of big data analytic in cardiovascular imaging can therefore be expected to be proportional to the capabilities and adaptabilities of information technology infrastructure of an organization and may require careful mandate in revamping the existing inefficiency.

Second, given the exponential trends in the growth of big data, the conventional pace at which the information technology infrastructure and decisions are modified will need to be changed; by the time there is some early outcome information available, the technology would have changed that would render the older information and decision outdated. Institutions will need to adopt a more continuous approach to analysis and decision making for real-time monitoring. The next-generation information technology processes with big data systems analytics would need to be designed for automated continuous real-time insights. This means that information technology applications need to measure and adapt simultaneously on a wide variety of issues, including cardiologist interactions, such as the appropriateness of imaging with the dynamic changes such as patient satisfaction and outcomes measured in real time, along with the changes in reimbursement policies. Such an integrated imaging architecture will essentially be an information ecosystem: a network of internal and external services continuously sharing information, optimizing decisions, communicating results, and generating new insights.

**Table 1.** Comparison of Traditional Versus Big Data Analytics

| | Traditional Data Analysis | Big Data Analysis |
|---|---|---|
| Personal | Statistician/ epidemiologist | Data scientist/technology entrepreneur |
| Participation | Solo | Works in a team |
| Inputs | Data file<br>Hypothesis driven | Enterprise-level information, problem driven |
| Data | Archived, siloed | Distributed, real time |
| Data type | Pre-prepared, clean | Messy, unstructured |
| Data size | Kilobytes | Gigabytes |
| Question focus | Conservative, inferential<br>What happened?<br>Why? | Risk taking, predicting<br>What will happen?<br>What if? |
| Tools | SAS, mainframe | R, Python, Hadoop, Linux, QlickView, Tableu, Jaspersoft, Saffron, etc. |
| Output | Tables, graphs | Advance visualization |
| Latency | Days to weeks | Seconds |
| Applicability | Historical, identifying confounding influences | Futuristic, correcting for confounding factors |

Adapted from Smith (8).

Furthermore, successful organizations will have to train and recruit a new breed of data scientists (Table 1) with novel sets of skills in big data visualization that can readily integrate multimodality and multidimensional datasets into the workflow (8). As data science emerges as a scientific discipline for personalized medicine, there are questions about the death of traditional population-based statistics. It is true that big datasets may have different properties than small datasets. In statistics, an expert individual drives the analysis following a predetermined strategy because the user has insufficient expertise to do so, whereas in large data analytics, the program would drive the analysis because the individual has insufficient resources to manually examine billions of records and hundreds of thousands of potential patterns. It is clear that one would have to move away from a conservative analysis to a more risk-taking attitude in automatic computing and disease modeling. This would also demand creating a cadre of imaging specialists who focus upon database processing algorithms for integrating new understanding into the knowledge domain.

In summary, imaging big data will prompt organizations to rethink their basic assumptions about the relationships among imagers, data scientists, and information technology professionals and their respective roles. Although the questions and potential solutions may be varied, there is little doubt about the manifold value that imaging big data will potentially provide, with new insights into disease, treatment, and interventions. This was essentially the purpose of highlighting an intuitive and futuristic viewpoint (2) by displaying the heat map on the cover page.

**Address for correspondence:** Dr. Jagat Narula, Icahn School of Medicine at Mount Sinai, Mount Sinai Heart, One Gustave L. Levy Place, Mailbox 1030, New York, New York 10029. *E-mail:* *jagat.narula@mountsinai.org*.

## REFERENCES

1. United States Patent and Trademark Office. Registration #75263259. September 1, 1993.
2. Sengupta PP. Intelligent platforms for disease assessment: novel approaches in functional echocardiography. J Am Coll Cardiol Img 2013;6:1206–11.
3. Larson EB. Building trust in the power of "big data" research to serve the public good. JAMA 2013;309:2443–4.
4. Murdoch TB, Detsky AS. The inevitable application of big data to health care. JAMA 2013;309:1351–2.
5. Kalil T, Zhao F. Unleashing the power of big data. Available at: http://www.whitehouse.gov/blog/2013/04/18/unleashing-power-big-data. Accessed September 17, 2013.
6. Poon EG, Jha AK, Christino M, et al. Assessing the level of healthcare information technology adoption in the United States: a snapshot. BMC Med Inform Decis Making 2006;6:1.
7. Harris Interactive. European physicians especially in Sweden, Netherlands and Demark, lead U.S. in use of electronic medical records. Health Care News 2002;2:1–3.
8. Smith D. Statistics vs data science vs BI. Available at: http://blog.revolutionanalytics.com/2013/05/statistics-vs-data-science-vs-bi.html. Accessed September 17, 2013.