

Editorial

Big data in radiation oncology

Radiation oncology (RO) is data rich as much as technology intensive. Decision-making in RO is algorithmic and is based on the evidence gathered over various platforms. A lot of data generated are in digital format such as in computed tomography (CT) scan, positron emission tomography, magnetic resonance imaging, sonography, on-board CT, and planning. There are already systems for archival and retrieving data.

The progress in the technology of radiation therapy has transformed from noisy Van-der Graffe generators to fully digital linear accelerators with various capabilities. Intensity-modulated radiation therapy, image-guided radiation therapy (IGRT), adaptive radiation therapy, biological planning, and deformable registration are the consequences of harnessing technology of the digital era. Humongous amount of digital data are created in the process of treatment of patients, which is in the nature of both dicom and nondicom formats. Decision-making of clinicians is generally algorithmic. Descriptive reporting of histopathology is getting synoptic while genomic data can have simple codes. The personal data of patients, diagnosis, and outcomes can be easily coded according to the standards. The outcomes such as disease-free survival, initial response, and overall survival are all fairly well defined and are included in the database. The nondicom images like that from endoscopy or clinical evaluation can be archived but it seems difficult to archive in a format which is easy to analyze. However, the conclusion arrived at from these images can be coded. The data so far have been archived and retrieved passively. Analysis based on accumulated data has been *post hoc*. Newer avenues of analytics and big data are changing it all.

Big data refers to computational analysis of voluminous structured or unstructured data which are in the range of terabytes to exabytes. The data are of no consequences without any analytic. It is relevant to make sense of the voluminous information such that hidden patterns, relations, and meanings emerge for

hypothesis generation and understanding. Big data is about volume, velocity, veracity, and variability. Big data research has application in health-care delivery as much as it is in weather forecasting or understanding trends in stock market. A big thrust for big data for RO started in 2013 by the National Institutes of Health, American Society for Radiation Oncology, and Association of Physicists in Medicine in the meeting, "Technology for innovations in Radiation Oncology." There has been a considerable thrust in developing big data for RO since then.

IGRT, commonly referred to as image-guided radiation therapy, can be redefined as information-guided radiation in the era of big data. Data can emerge from structured and nonstructured electronic health records, diagnostic images, on-board images, information on genomics besides histology, and patients' outcomes. Veracity of data is crucial for effective and reliable analytics. It is a big gap that needs to be filled before meaningful patterns can emerge. Variability of language and inherent in homogeneity is another road block to be surmounted. An effort to ethical consideration for big data is evolving. Data archival and retrieval has to be Health Insurance Portability and Accountability Act compliant. De-identified data sets do not need any informed consent. The ethical concerns of big data are getting apparent. However, staffs who handle voluminous data are alien to the idea of ethics. It is mandatory to sensitize everyone involved in the issues related to confidentiality. Misrepresentation of data, poor understanding of statistical and scientific methods, and disregard to accuracy of data collection and collation can affect the outcomes of any analysis. Data have to be protected from hackers.

Computerized treatment planning with inverse planning is de rigueur now. Library-based segmentation is evolving as much as

Nagraj Huilgol

Department of
Radiation Oncology,
Nanavati Super
Speciality Hospital,
Mumbai, Maharashtra,
India

For correspondence:

Dr. Nagraj Huilgol,
Department of
Radiation Oncology,
Nanavati Super
Speciality Hospital,
Mumbai,
Maharashtra, India.
E-mail: nagrajhuilgol@gmail.com

Access this article online

Website: www.cancerjournal.net

DOI: 10.4103/0973-1482.194598

Quick Response Code:



Cite this article as: Huilgol N. Big data in radiation oncology. J Can Res Ther 2016;12:1107-8.

This is an open access article distributed under the terms of the Creative Commons Attribution-NonCommercial-ShareAlike 3.0 License, which allows others to remix, tweak, and build upon the work non-commercially, as long as the author is credited and the new creations are licensed under the identical terms.

Huilgol: Big data

knowledge-based radiation therapy. Watson backed by big data and running on machine language is getting as smart as human-based decision-making. Artificial intelligence (AI) is going to grow and permeate the health-care practice. It can be a blessing for underdeveloped countries with

low human resources or those countries with dwindling population. Big data in health care is in its inception but has the potential to be a disruptive innovation. Big data and AI do come with inherent flaws and risks. It is time to march on with eyes open.

