Journal of the Royal Society of Medicine; 2019, Vol. 112(12) 493–494
DOI: 10.1177/0141076819871055

Data science in modern evidence-based medicine

Dina Radenkovic[1], Sir Bruce Keogh[2] and Mahiben Maruthappu[3]
[1]Guy's & St Thomas' NHS Foundation Trust, Westminster Bridge Road, London SE1 7EH, UK
[2]Birmingham Women's and Children's NHS Foundation Trust, Mindelsohn Way, Edgbaston, Birmingham B15 2TG, UK
[3]Cera Care, Elizabeth House, 4th, London SE1 7NQ, UK
Corresponding author: Dina Radenkovic. Email: dina.radenkovic@gmail.com

To practise evidence-based medicine, clinicians need to apply the findings of scientific research to the circumstances of individual patients as part of their decision-making process. For centuries, medicine and clinical reasoning was based on subjective personal experience, until the wider adoption of evidence-based medicine, defined as '*The conscientious explicit and judicious use of current best evidence in making decisions about the care of individual patients*'.[1]

To date, simple mathematical and statistical methods have been used to describe patterns within relatively small-size datasets. We are now faced with a generation of novel datasets in large quantities, from an almost infinite number of digital fingerprints (phenomes) and sources such as electronic health records, high-resolution medical imaging, wearable devices and biosensors, more widely available genetic testing and even inputs from social media.[2] It appears that the standard means of critical appraisal and data interpretation have reached a point of saturation. Evidence-based medicine, as we know it today, is necessary, but not sufficient to meet the demands of analysing large and complex datasets, as it can be time-consuming, resource-intensive, slow to completion and expensive.

Data science is still a novel field, at the intersection of several disciplines including computer science, mathematics, statistics, and health and business analytics. It has evolved across many industries to be an advanced statistical method of gathering insights from large amounts of data in order to achieve business objectives or understand user behaviours. We are at a point in healthcare where data are more widely available and accessible in a computer readable fashion than before, coincided with the surge in computer processing power that occurred over the past decade.[3]

Recently, there has been significant interest in the field of artificial intelligence and its applications to healthcare and artificial intelligence was one of the major topics of the recently published Topol Review.[4] Artificial intelligence broadly implies the operation of computer programmes with human cognitive ability in order to enable automated solving of complex problems, including perception, recognition, memory and learning. With a growing number of massive datasets and the increased computational power of the machines, a more unique type of artificial intelligence called machine learning has emerged. Machine learning transforms the inputs of an algorithm into outputs using statistical, data-driven rules that are automatically derived from a large set of examples, rather than being explicitly specified by humans.[4] More specifically, a subfield of machine learning called deep learning where the algorithms are created from the data itself and determined by the number of layers (unsupervised learning) rather than initial human rules and postulates (supervised learning) is gaining increasing application in healthcare and already showing to compare with clinicians in the field of imaging, pathology, skin disease diagnosis, ophthalmology, cardiac arrhythmia detection and even sepsis management.[5] Although the majority of the described examples have not been tested in pragmatic clinical trials, they hold significant potential, primarily because they could enable assessment, monitoring, diagnosis and management of a greater number of patients at a lower cost.[3] In addition, artificial intelligence has the potential to reduce diagnostic errors as a quality assurance tool, as well as being used in workflow management. Consequently, it could reduce the administrative burden of clinicians and provide them with the 'gift of time' to prioritise treating patients.[6]

To derive meaningful inferences from newly generated data, the data need to be stored and analysed with advanced statistical methods; it appears that the next step of evidence-based medicine could be data science with processing software through increased input of machines. Therefore, different data analytics skills, particularly basic familiarity of machine learning concepts, are becoming important for clinician scientists. Nevertheless, the core principles of evidence-based medicine are still fundamental in the era of big data. Data science innovations must be evaluated with the same epistemological rigor that lies at the heart of evidence-based medicine. On the other hand, the standard approach to evidence-based medicine needs the addition of machine learning as a critical tool in order to continue to enable clinical decision-making based on the synthesis of all available evidence.

Qualifying and working in medicine at this time where big data and machine learning are suddenly opening a new horizon is exciting but care should be taken to avoid overwhelming the clinicians. Even before the described wave of digitalisation, medicine has been one of the most information-heavy and demanding professions. Training doctors may find it difficult to stay up to date with their medical knowledge and simultaneously develop novel skillsets. Thus, despite the challenges of implementation into formal medical undergraduate and postgraduate training, the responsibility to educate the new cadre of medical doctors equipped with higher digital literacy and appreciation of data science lies with the universities and board examiners.

Data science is a rapidly evolving multidisciplinary field and the regulatory bodies must keep up the pace to provide standards for novel research models, innovation and trends in data analytics. However, it is certain that the greater knowledge of computer programming by doctors is required for highly efficient, truly data-driven modern medicine.

Declarations

Competing interests: None declared.

Funding: None declared.

Ethics approval: Not applicable.

Guarantor: DR.

Contributorship: DR and MM had the idea for the article. All authors contributed to writing the manuscript and approved the final version. The corresponding author attests that all listed authors meet authorship criteria and that no others meeting the criteria have been omitted.

Acknowledgements: None.

Provenance: Not commissioned; editorial review.

ORCID iD: Dina Radenkovic https://orcid.org/0000-0003-1355-6967

References

1. Sackett DL, Rosenberg Wm, Gray JA, Haynes RB and Richardson WS. Evidence based medicine: what it is and what it isn't. *BMJ* 1996; 312: 71–72.
2. Hemingway H, Asselbergs FW, Danesh J, Dobson R, Maniadakis N, Maggioni A, et al. Big data from electronic health records for early and late translational cardiovascular research: challenges and potential. *Eur Heart J* 2018; 39: 1481–1495.
3. Topol EJ. High-performance medicine: the convergence of human and artificial intelligence. *Nat Med* 2019; 25: 44–45.
4. Topol E and Committee. Preparing the healthcare workforce to deliver the digital future. Independent report on behalf of the Secretary of State for Health, 2019. See https://topol.hee.nhs.uk/.
5. Esteva A, Robicquet A, Ramsundar B, Kuleshov V, DePristo M, Chou K, et al. A guide to deep learning in healthcare. *Nat Med* 2019; 25: 44–56.
6. Joshi I. Waiting for deep medicine. *Lancet* 2019; 393: 1193–1194.