IJHS

Editorial

Should we teach computational thinking and big data principles to medical students?

Mohammed Saqr[1,2], Matti Tedre[1]

[1]School of Computing, University of Eastern Finland, Joensuu Campus, Yliopistokatu, Joensuu, Finland, [2]Department of Computer and System Sciences (DSV), Stockholm University, Borgarfjordsgatan, Kista, Sweden

Address for correspondence:
Mohammed Saqr
School of Computing, University of Eastern Finland,
Joensuu Campus, Yliopistokatu 2, P.O. Box 111,
fi-80100 Joensuu, Finland

WEBSITE: ijhs.org.sa
ISSN: 1658-3639
PUBLISHER: Qassim University

The sources, nature, and scale of data are changing, and computing is deeply affecting the very nature of research from exploration, hypothesis creation, and literature review to collection, generation, and analysis of data, as well as reporting of results. In many fields, no large-scale experiments are funded unless computer simulation shows high likelihood for results to turn out as expected. What is more, sensor technology and communication technology are making information technology ubiquitous, embedded in all sectors of life. As health-care practitioners to have become generators of data, users of data-driven decisions, and researchers using data, there is a need to rethink our curricula. Medical curricula should embrace the fundamentals of big data, the info-computational paradigm of science, modern inferential techniques, and algorithmic thinking principles. By understanding such principles, health-care practitioners will have improved skills for problem-solving in computerized research, they will better understand the insights generated by the devices they use, will be able to design better research, and most importantly will deliver a better service to patients. That understanding is also important from the perspective of bottom-up innovation: The better health-care practitioners understand the potential and caveats of computational thinking and big data, the better they are able to suggest new and improved ways of collecting data and dealing with it.

Electronic health records (EHRs) encompass a wide range of data about patients that include demographics, health condition, drug history, and investigations records along with financial and occupation history and more. EHRs were designed to be widely shared across institutions and governmental organizations. Several countries have national EHR initiatives that collect data about all citizens.[1] The growing adoption of EHR, along with the rapid digitization of medical devices, mobile phones and their rich sensor ecosystems, wearables, social media, and internet, has significantly changed the repertoire of data available about patients. The ability to have high resolution, high-dimensional data about large numbers of patients, sometimes in the real time.[2]

As data has grown exponentially in volume and resolution, so has computing power, storage capacity, and processing speeds. In addition to the development of hardware, development of algorithms, optimization techniques, development environments, simulation, and data management constitute roughly half of the speedup of scientific computing in the past 70 years. A burgeoning data science discipline has evolved in applications and accomplishments, technologies such as machine learning, and improved understanding of how and where new computational models can bring real solutions to challenging life problems such as optimum drug dosage, detection of pathologies, analysis of X-ray images, and assisting medical decision-making. The bounty of such technologies has stimulated companies to incorporate artificial intelligence (AI) insights into everyday medical devices used by physicians.[3,4]

While biology and bioinformatics are among the most cited examples of the info-computational paradigm shift in science,[5] health care is not exempt of the paradigm shift in how data are generated, analyzed, and interpreted. Health-care practitioners are generators of data about patients, use data in research about patients, and use data to guide their work. Most importantly, health-care practitioners use insights from digital medical devices, software, and computer-based tools every day. For example, an AI-enabled magnetic resonance imaging machine can help a radiologist identify and diagnose a brain tumor. An AI-enabled ECG can warn a cardiologist that the patient has a high likelihood of myocardial infarction. An intensive care unit doctor may get recommendations from an AI-enhanced ventilator of how to optimize the ventilation mode or settings. How these devices make such decisions and how accurate these decisions are is in the heart of medical care, as doctors use these insights into make countless decisions every day. Furthermore, patients are increasingly using their own AI-based devices to monitor their own health and activity through, for instance, heart rate monitors and tremor analyzers. Furthermore, the evolution of big data has brought forth new ethical and privacy challenges, those privacy concerns face practitioners in their everyday practice. As such they need to understand and handle such ethical issues.[2,3,6,7]

Data-driven research provides alternative approaches to older types of empiricism. Unlike theory-driven traditional research and unlike older, manual exploration of data sets, big data analytics aims at using massive amounts of data with inductive reasoning to generate insights, identify weak signals, and unearth phenomena not visible in smaller data sets. The data sets used in the era of big data are intractable for humans to analyze and yield only to massive computing power. Data scientists usually try to explore massive, high resolution, high-dimensional data with no prior hypothesis to find interesting patterns and correlations. The premise is that such bottom-up approach may lead to the discovery of realities we did not know exist. There are many instances that support this mode of discovery, but also instances that have led to spurious correlations.[8]

Given the emergence of a new, info-computational paradigm of science as well as numerous computational and data-driven tools for work in our fields, there is a need to rethink our curricula. Medical curricula should embrace the fundamentals of big data, modern inferential techniques, and principles of computational thinking. To grow familiar with the 2000s scientific techniques and technology, students will need to learn to see and interpret the world through computational and data-driven lenses. By understanding the principles involved in computational sciences, health-care practitioners will have improved problem-solving skills well suited for the latest technology, will better understand the insights generated by the devices they use, will design better research, and, most importantly, will deliver a better service to patients. However, this is not a call for teaching mathematics or programming; it is rather a call for improved data literacy, fluency with information processes and streams, and computational thinking. Those are focused on understanding what data streams are available, how algorithms and machine learning use and process data, how dumb machinery makes smart inferences and decisions, what computers can and cannot do, and how to best combine strengths of humans and computers and mitigate the weaknesses of them both.[9,5] Computational thinking is the best taught integrated in courses where computational tools and methods are used in real work setting but can also be integrated in research methodology courses where principles of big data methods, machine learning, and modern inference can be incorporated. Computational thinking can be discussed within the framework of interdisciplinary problem-solving, basic sciences, or faculty development seminars.

The doors should be open to innovation: Medical students spend a great deal of their education studying basic sciences, and much of that can be interwoven into everyday practice with data and computers. Medicine should be responsive, integrative, and interdisciplinary. Educating the health-care personnel of the future must remain open to current innovations or we may be risking graduating doctors trained with the thinking tools for the past 30 years instead of the thinking tools for the next 30 years.

References

1. Häyrinen K, Saranto K, Nykänen P. Definition, structure, content, use and impacts of electronic health records: A review of the research literature. Int J Med Inform 2008;77:291-304.
2. Saqr M. Big data and the emerging ethical challenges. Int J Health Sci (Qassim) 2017;11:1-2.
3. Topol EJ. High-performance medicine: The convergence of human and artificial intelligence. Nat Med 2019;25:44-56.
4. Chen H, Chiang RH, Storey VC. Business intelligence and analytics: From big data to big impact. MIS Q 2012;36:1165-88.
5. Denning PJ, Tedre M. Computational Thinking. Cambridge, MA, USA: The MIT Press; 2019.
6. Schneeweiss S. Learning from big health care data. N Engl J Med 2014;370:2161-3.
7. Saqr M. Beyond the sophomoric promises of artificial intelligence in medicine. Int J Health Sci (Qassim) 2018;12:1-2.
8. Kitchin R. Big data, new epistemologies and paradigm shifts. Big Data Soc 2014;1:1-12.
9. Tedre M, Denning PJ. The Long Quest for Computational Thinking. Proceedings of the 16th Koli Calling International Conference on Computing Education Research. New York, USA: ACM; 2016. p. 120-9.