Canadian Journal of Cardiology 36 (2020) 19–21

Editorial

The Role of Physicians in the Era of Big Data

George A. Heckman, MD, MSc, FRPCP,[a,b] John P. Hirdes, PhD,[a] and Robert S. McKelvie, MD, PhD, FRCPC[c]

[a] *School of Public Health and Health Systems, University of Waterloo, Waterloo, Ontario, Canada*

[b] *Schlegel-UW Research Institute for Aging, University of Waterloo, Waterloo, Ontario, Canada*

[c] *Schulich School of Medicine and Dentistry, Western University, London, Ontario, Canada*

See article by Buchan et al., pages 84–91 of this issue.

HAL, despite your enormous intellect, are you ever frustrated by your dependence on people to carry out your actions?[1]

The core business of medicine, succinctly summarized by Dhaliwal and Detsky,[2] is to answer 3 patient questions: 1) What is happening to me? (diagnosis); 2) What will become of me? (prognosis); and 3) What can be done about it? (management). This business is centuries old. More than 3500 years ago, Egyptian physicians could identify the clinical syndrome of angina, and though their ability to manage it was limited, they understood its prognosis: "If thou examinest a man for illness in his cardia and he has pains in his arms, and in his breast and in one side of his cardia, it is death threatening him."[3]

Our ability to answer these questions made great strides over the past several centuries, with advances not only in the understanding and treatment of disease, but also in broader determinants of health. In addition to gains in socioeconomic status, the latter half of the 20th century heralded great improvements in public health and the treatment of conditions that before had severely limited prognoses. As a result, many nations are seeing their populations age. Although that is an achievement to celebrate, it comes at a cost: People who live longer do so with a rising burden of chronic disease and age-associated cognitive impairment, disability, and frailty.[4,5]

Patients today, particularly those who are aging with chronic diseases, are complex. Treatment options abound, care is increasingly interdisciplinary, and the health care system is more byzantine. Growth in information is exponential: Improvements in diagnostic testing, the emerging concepts of frailty and resilience, innovations in biomarkers, genomics, and precision medicine, and the advent of wearable monitors will drive a rapid expansion in the amount of data about individual patients.[6-8] Paradoxically, this "big data" explosion may be overwhelming our ability to answer patients' questions.

Complexity of this degree can not only exceed the cognitive capacity of physicians and other providers, leading to a higher risk of error, but failure to provide answers to patient questions may also interfere with patient engagement in self-care and lead to more harm.[2,9] Clinical decision-support tools are intended to assist clinicians and patients by synthesizing health information and providing, for example, estimates of risk and prognosis. Typically developed with the use of statistical modelling, these tools are now increasingly developed with the use of machine learning and artificial intelligence (AI) techniques. How well do clinical decision-support tools perform compared with physicians, assuming that such tools are even used?

In this issue of the *Canadian Journal of Cardiology*, Buchan et al. present results from a cross-sectional study of 150 patients attending a tertiary care heart failure management program, and in which they compared prognoses established with the use of 3 different validated instruments and those estimated by attending cardiologists and family physicians.[10] Scores predicted by means of the HF Meta-Score, Seattle Heart Failure Model, and Meta-Analysis Global Group in Chronic HF scores correlated well with one another. In contrast, agreement was limited between the physicians and the instruments as well as among the physicians, with cardiologists offering more pessimistic prognoses than family physicians. Although the study design precluded comparison of these estimates with actual prognoses, the authors point to other data suggesting that physicians tend to overestimate mortality compared with validated instruments. Very few of the physicians in this study reported routinely using such instruments in practice. These results raise several questions, not just related to physician training. In an era of AI, how can clinical decision-support tools be embedded into routine practice? How should the physician's role adapt to more effectively answer the 3 patient questions?

Received for publication September 5, 2019. Accepted September 26, 2019.

Corresponding author: Dr George A. Heckman, Schlegel Research Chair in Geriatric Medicine, Schlegel-UW Research Institute for Aging, Associate Professor, School of Public Health and Health Systems, University of Waterloo, LHN 3706, 200 University Avenue West, Waterloo, Ontario N2L 3G1, Canada. Tel.: +1-519-888-4567-31028.

E-mail: ggheckma@uwaterloo.ca

See page 20 for disclosure information.

https://doi.org/10.1016/j.cjca.2019.09.018

Clinical judgment denotes the cognitive tasks that physicians apply to answer patient questions.[11] Clinical judgment rests on a dual foundation of heuristics and intuition honed through experience and of formal knowledge gained through the application of the scientific method to clinical assessment.[12] The application of the scientific method to formalize clinical knowledge is the raison d'être of evidence-based medicine: to improve the reproducibility of diagnoses, prognoses, and treatment recommendations. However, this approach is inherently population based, and the application of generalizable data to individual patients remains challenging.[11] The findings of Buchan et al., notably the differences between specialists and family physicians regarding prognoses offered, suggest that differences in clinical experience during training and practice may affect the development of clinical judgment heuristics. Moreover, while the clinical decision-support tools used in this study apply to a specific patient population, they are not readily applicable to the broader population of heart failure patients, who are generally older, more often women with preserved ejection fraction, and with various degrees of multimorbidity and frailty.[13] This is potentially problematic, because we know that in the face of increased patient complexity, the risk of clinical judgment errors rises.[2]

The promise of AI lies in its ability to more fully capture patient complexity by incorporating increasingly large amounts of data into clinical decision-support tools and thereby improve their diagnostic and predictive abilities.[11,14,15] AI is now being applied to all areas of medicine for data analysis, to improve diagnosis, and to support care planning. However, clinicians need to be aware that pitfalls remain. First, a fundamental problem lies in the quality of the data used to elaborate a clinical decision support tool. For a prognosis to be accurate, the data on which it is based must be reliable and valid to avoid "garbage in, garbage out." Data must also be complete and comprehensive, and fully represent all pertinent characteristics of a population, such as gender, sex, body habitus, ethnicity, geography, socioeconomic status, frailty, and resilience, among many other factors. Experience so far suggests that AI-based systems can still make glaring mistakes.[16] Second, AI-derived tools are typically evaluated on their ability to correctly classify individuals according to the occurrence of a certain event (eg, death) with the use of metrics such as the area under the receiver operating characteristic curve or the C-statistic.[11] Higher model fit does not necessarily imply that, at the level of individual patients, the fit is perfect nor that confidence intervals around point estimates have been eliminated.[11-15] Estimates of risk still reflect group averages, even if the groups are somewhat smaller. Third, even the best instrument is of no help if it does not inform outcomes that matter most to patients, which is particularly relevant to those nearing the end of life. Fourth, secular trends have been shown to be a significant source of variability in prognostic models that, regardless of the method used to derive them, are based on past data and thus cannot easily account for emerging factors that could mediate risk.[15] Finally, whereas traditional prognostic algorithms are derived based on well understood theories or basic mechanisms, machine learning and AI techniques search for complex associations in large data sets and may arrive at solutions that are difficult for clinicians to understand, let alone explain to patients.[11,14,17,18]

There remains an important aspect of the clinician's task that AI is unlikely to replace. The application of clinical judgment is not a unilateral process wherein the physician simply relays information. It is to engage patients in a conversation. True person-centered care can be achieved only through dialogue, which requires not only sharing of the most accurate information possible, but that patients be supported to interpret "the science" within the context of their personhood, their goals and their preferences.[11,19] We have all known patients with similar health problems who make diametrically opposed therapeutic choices based on personal preferences.

Clinicians must understand that for AI to be clinically useful, it requires robust, reliable, and valid data standards across the health care system. They need to expect and demand that clinical decision-support tools based on AI undergo rigorous psychometric testing before being broadly implemented. Clinicians still require a robust understanding of the epidemiology and mechanisms of disease, particular when the logic that underlies an AI-derived solution is not readily evident, to ascertain that all relevant data elements were included in derivation data sets. This is particularly crucial with the democratization of AI. Access by patients to their own health information, wearables, and online AI systems has great potential to support better patient outcomes, but physicians require education and vigilance to identify and assist those at risk of harm from the misapplication of such systems. Electronic medical records require significant upgrades to support high data quality and facilitate the use by clinicians of clinical decision-support systems.[20] Finally, clinical experience remains an essential component of clinical judgment, and the extent to which different training environments may have caused Buchan et al. to observe more pessimistic estimates of prognosis by cardiologists compared with family doctors needs to be explored. Perhaps specialty trainees should spend more time in the community with primary care providers.

The adoption of AI in medicine will not eliminate the need for clinicians. Indeed, the importance of highly trained physicians, with superlative communication skills, will only grow. To continue answering the 3 patient questions, physicians must become proficient at exploiting the value of AI, and recognize and contend with its inherent limitations. The future is now.

Funding Sources

George A. Heckman receives salary support from the Schlegel Research Chair in Geriatric Medicine.

Disclosures

The authors have no conflicts of interest to disclose.

References

1. Kubrick S. director. 2001: A Space Odyssey. Stanley Kubrick Productions, 1968.

2. Dhaliwal G, Detsky AS. The evolution of the master diagnostician. JAMA 2013;310:579-80.

3. Saba MM, Ventura HO, Saleh M, Mehra MR. Ancient Egyptian medicine and the concept of heart failure. J Card Fail 2006;12:416-21.

4. Strong K, Mathers C, Leeder S, Beaglehole R. Preventing chronic diseases: how many lives can we save? Lancet 2005;366:1578-82.

5. Junius-Walker U, Onder G, Soleymani D, et al. The essence of frailty: a systematic review and qualitative synthesis on frailty concepts and definitions. Eur J Intern Med 2018;56:3-10.

6. Mitnitski A, Rockwood K. Aging as a process of deficit accumulation: its utility and origin. Interdiscip Top Gerontol 2015;40:85-98.

7. Hale M, Shah S, Clegg A. Frailty, inequality and resilience. Clin Med (Lond) 2019;19:219-23.

8. Beckmann JS, Lew D. Reconciling evidence-based medicine and precision medicine in the era of big data: challenges and opportunities. Genome Med 2016;8:134.

9. Riegel B, Moser DK, Anker SD, et al. State of the science: promoting self-care in persons with heart failure: a scientific statement from the American Heart Association. Circulation 2009;120:1141-63.

10. Buchan TA, Ross HJ, McDonald M, et al. Physician judgment vs model predicted prognosis in patients with heart failure. Can J Cardiol 2019;36:84-91.

11. Chin-Yee B, Upshur R. Clinical judgement in the era of big data and predictive analytics. J Eval Clin Pract 2018;24:638-45.

12. Norman G. Dual processing and diagnostic errors. Adv in Health Sci Edu 2009;14:37-49.

13. Heckman GA, McKelvie RS, Rockwood K. Individualizing the care of older heart failure patients. Curr Opin Cardiol 2018;33:208-16.

14. Liu X, Keane PA, Denniston AK. Time to regenerate: the doctor in the age of artificial intelligence. J R Soc Med 2018;111:113-6.

15. Pate A, Emsley R, Ashcroft DM, Brown B, van Staa T. The uncertainty with using risk prediction models for individual decision making: an exemplar cohort study examining the prediction of cardiovascular disease in English primary care. BMC Med 2019;17:134.

16. McDonald G. Danger, danger! 10 alarming examples of AI gone wild. InfoWorld. March 23. Available at: https://www.infoworld.com/article/3184205/danger-danger-10-alarming-examples-of-ai-gone-wild.html. Accessed September 28, 2019.

17. Fischer T, Brothers KB, Erdmann P, Langanke M. Clinical decision-making and secondary findings in systems medicine. BMC Med Ethics 2016;17:32.

18. London AJ. Artificial intelligence and black-box medical decisions: accuracy versus explainability. Hastings Cent Rep 2019;49:15-21.

19. Chin-Yee B, Messinger A, Young LT. Three visions of doctoring: a gadamerian dialogue. Adv Health Sci Educ Theory Pract 2019;24:403-12.

20. Howe JL, Adams KT, Hettinger AZ, Ratwani RM. Electronic health record usability issues and potential contribution to patient harm. JAMA 2018;319:1276-8.