Clinical considerations when applying machine learning to decision-support tasks versus automation

Trevor Jamieson,[1,2,3] Avi Goldfarb[4]

[1]Department of Medicine, University of Toronto, Toronto, Ontario, Canada
[2]WCH Institute for Health System Solutions and Virtual Care (WIHV), Women's College Hospital, Toronto, Ontario, Canada
[3]Division of General Internal Medicine, St Michael's Hospital/Unity Health Toronto, Toronto, Ontario, Canada
[4]Rotman School of Management, University of Toronto, Toronto, Ontario, Canada

Correspondence to
Dr Trevor Jamieson, Division of General Internal Medicine, St Michael's Hospital, Toronto M5B 1W8, Canada; jamiesont@smh.ca

[Linked]

► http://dx.doi.org/10.1136/bmjqs-2018-008977

Check for updates

The future role of clinical automation in healthcare is a matter of debate, from commenters who claim that artificially intelligent clinical entities could relatively easily replace 80% of what physicians do[1] to those who see a future of a "well-informed, empathetic clinician armed with good predictive tools and unburdened from clerical drudgery".[2] While the extent to which clinicians will be able to be replaced by machines is a larger topic than will be covered here, what is clear is that artificial intelligence will transform the way healthcare is delivered.[3 4]

In this issue of *BMJ Quality and Safety*, for example, we see a report on a randomised controlled trial (RCT) of the use of a robot to capture historical information from older adults.[5] Boumans *et al* randomised 42 community-dwelling seniors to have a 52-item questionnaire captured by a nurse or a social robot, allowing for the generation of three indices of frailty, well-being and resilience. In this small pilot, the robot completed the vast majority of interviews without assistance (92.8%) and the interview time and index scores were comparable, although it would be incorrect to suggest that the performance was interchangeable. The robot interviews showed much less variation in duration. Nurse interviews lasted an average of 15 min but with a wide SD of 8.5 min. The robot interviews lasted an average of 16.6 min (p=0.2 for comparison with nurse interviews) but with a SD of only 1.5 min. In other words, assigning these interviews to a robot would result in a much more predictable time commitment for patients.

In their Discussion, Boumans and colleagues write that because "Many

people are concerned about robots taking over human jobs…", it is more palatable to introduce the robot as an assistant rather than as a replacement. Nonetheless, many observers will clearly regard the primary justification for the robot as freeing the nurse from a time-consuming task or, stated another way, replacing the human performing a task with a robot. From the perspective of health quality, it remains unclear whether the optimal future state for any given task will be one of human superiority, machine superiority or a synergistic partnership that is greater than the sum of its parts—in the specific case of computer-assisted mammography, there is a suggestion that the latter could be true.[6] Rather than dwell on what remains a largely philosophical question at this point in time, we elected to use the opportunity afforded by the study of Boumans *et al* to highlight some of the important clinical considerations for artificially intelligent systems that serve to support, and ideally augment, rather than to replace.

The recent attention to artificial intelligence has been driven by advances in a particular subfield of computer science called machine learning. Machine learning, a form of computational statistics, is based on algorithms that use data to generate predictions. These predictions—defined as the process of filling in missing information—allow machines to perform tasks without explicit instructions and can be combined with other algorithms to enable either *automation* or *decision support*.[7] In automation, a machine operates independently to complete a task, whereas in decision support, a machine is concerned with providing information

or assistance to the primary agent responsible for task completion. In the included RCT, under automation, the robot would complete the historical task completely independently, whereas under decision support, the robot might capture a history to the best of its ability and then provide that information to the nurse who would then use that information to confirm, augment or even simply approve of the captured information. With decision support, clinical decisions rest with the clinicians and depends on their individual judgements of the consequences of different actions.

We expect machine learning to lead to automation when (1) a human prediction takes time and effort, (2) when human judgement can determine what to do with a prediction long before the prediction is made, for example when there is little need for personalisation, (3) and when the workflows of the healthcare practitioner are unlikely to change. Automation is already happening in radiology, but not in the dramatic ways that casual observers might expect. Automating the interpretation of radiological images would require intelligent technology and also substantial modifications to the radiologists' workflows and a non-trivial shifting of accountability resulting in regulatory and practical barriers. This is not where we are seeing the immediate shift. Proving the point that automation is more likely when the surrounding workflows are minimally impacted, the actual impact in radiology has been in documentation—a key bottleneck in a radiologist's workflow. Until recently, human transcriptionists translated audio recordings into formatted text, but increasingly the transcriptionist is replaced by a machine that automatically turns the voice recordings into typed notes allowing for real-time, rather than delayed, confirmation and modification. In this case, automation is relatively straightforward because the radiologist's workflow becomes more efficient but does not change substantially.

The distinction between automation and decision support is critical—when deploying such a system, clarity on whether the goal of the endeavour is to automate the activity, that is, to replace the human component, or to provide decision support to the activity, that is, to augment the human component, has major consequences. While it may be assumed that decision support is simply a stepping stone on the progression towards full automation, the truth is that decision-support systems have fundamentally different considerations that must be accounted for in design and implementation. Specifically, those implementing artificially intelligent systems with an eye to providing decision support (vs automation) must be clear on the nature of the support and how it is integrated into other tasks, how trust of that support is established and how labour may be, or is desired to be, impacted.

First, in terms of nature of the support provided, doctors are already tasked with making complex decisions in a complex system,[8] using inefficient tools that

may be contributing to burnout,[9][10] and in an environment filled with interruptions.[11–13] While decision support could provide a much needed reprieve, if poorly integrated into a system it could also significantly increase workload—by increasing the volume of data entry in order to generate useful predictions, and cognitive load—by providing those predictions without a view to the cognitive effort to process and use the information.[14] Even in binary decisions, there is ample evidence that physicians, even those with dedicated statistical training, have poor comprehension of basic statistical measures relevant to healthcare decisions,[15] and greater computational power opens the door to much more complex non-binary decisions and the overlay of choice overload.[16][17]

A key supporting technology of decision-support systems will likely, therefore, be data visualisation.[18] In computer-assisted mammography, for example, the computer annotates the images to draw the human's attention to problem spots; this is entirely different from providing a list of problematic pixels—even if the data are identical. It is notable that two recent articles on quality and safety issues with artificial intelligence in this very journal paid only scant reference to the human–machine interface as a critical component of the artificially intelligent decision-making apparatus.[19][20] If the goal is automation, these are non-issues, but if the goal is decision support, the questions of the workload involved in getting the algorithms the data they need and the interpretability of the results for time-constrained decision-makers are critical success factors.

Second is the question of trust. The clinicians need to trust the guidance provided by the machine, and then (transitively) the clinicians must be able to translate that trust into a shared decision-making process with the patient. In 1995, at the advent of an explosion in the use of clinical epidemiological techniques to generate prognostic models, clinical credibility of a model was felt to require that a "model's structure should be apparent and its predictions should make sense to the doctors who will rely on them".[21] This requirement is obviously a problem with the typical deep learning 'black box', and the need for algorithmic transparency in domains such as health and law has led to an entirely new field of 'explainable AI'.[22] It may not always be true that explanation is inherently required when using a machine prediction to support a decision; if the predictions are accurate and lead to better outcomes as evidenced through the rigour of controlled investigation, it likely will not limit clinician acceptance any more than a lack of a detailed understanding of a biochemical mechanism limits their prescribing of a pharmaceutical. The challenge will be in situations where that evidence does not come or in situations where, despite rigorous evidence, the algorithms are hindered by generally poor data availability and quality leading to reduced trust through the assumption of 'garbage in, garbage out'.[23]

Regardless, decision support will not work without trust, and designers of such decision-support systems must build them with careful consideration of how that trust might be established.

Third, while automation has a relatively straightforward impact on the labour of the person whose job is automated, the impact on labour in a decision-supported system can be subtler and requires careful consideration. Decision-support systems can increase system efficiency primarily by increasing throughput, with variable impact on costs depending if labour costs are fixed and capitated versus fee for service. Other decision-support systems may achieve efficiencies more through a process of de-skilling. With de-skilling, the decision support allows a task to be completed with reduced training and expertise, thus allowing a shifting of tasks to lower-paid professionals, like nurses and pharmacists. Certainly, there would be regulatory barriers to this process, but it is already occurring in other contexts and high-quality decision support would make this process easier. In other circumstances, people may envision a relatively neutral impact on throughput, with no de-skilling, but rather that the decision support would free the medical professional from time-consuming administrative tasks, thus allowing them to engage in the oft marginalised humanist 'art of medicine'.[2 24] While throughput and de-skilling have more concrete traditional economic impacts, the impact of engagement in the art of medicine is highly indirect and thus may require more management to achieve.

In any event, it is key for designers and implementers of decision-support systems to have an understanding of what the envisioned labour impact of the system is, as that will determine the optimal nature of support and to whom. One must also consider whether existing regulations or the nature of the existing workforce, for example, unionised or not, will make the desired impact on labour, efficiency and decision-making impossible.

In summary, while recent advances in artificial intelligence will sometimes lead to automation, many applications in medicine will ultimately relate to decision support. Such decision support should not be seen as 'automation lite'. Decision support is different. It requires careful attention to the human–machine interface, specifically the nature of the support and its informational complexity, and the establishment of trust. Furthermore, it will affect labour by enabling either more efficient decisions, more human-to-human interaction or both. Implementing new systems in healthcare requires a clear vision of what you are trying to accomplish. Well-designed decision-support systems will facilitate workflows and decision-making, enable trust and more optimally leverage the human component of systems. We believe these design efforts will ultimately pay off by allowing higher quality and more efficient care.

REFERENCES

1 Khosla V. "20 percent doctor included" & Dr. Algorithm: speculations and musing of a technology optimist, 2016. khoslaventures.com[updated 2016 Sep 30; cited 2019 May 15]. Available: https://www.khoslaventures.com/20-percent-doctor-included-speculations-and-musings-of-a-technology-optimist

2 Verghese A, Shah NH, Harrington RA. What this computer needs is a physician: humanism and artificial intelligence. *JAMA* 2018;319:19–20.

3 Naylor CD. On the prospects for a (deep) learning health care system. *JAMA* 2018;320:1099–100.

4 Hinton G. Deep learning—a technology with the potential to transform health care. *JAMA* 2018;320:1101–2.

5 Boumans R, van Meulen F, Hindriks K, *et al*. Robot for health data acquisition among older adults: a pilot randomised controlled cross-over trial. *BMJ Qual Saf* 2019;28:793–9.

6 Rodríguez-Ruiz A, Krupinski E, Mordang J-J, *et al*. Detection of breast cancer with mammography: effect of an artificial intelligence support system. *Radiology* 2019;290:305–14.

7 Agrawal A, Gans JS, Goldfarb A. Artificial intelligence: the ambiguous labor market impact of automating prediction. *J Econ Perspect* 2019;33:31–50.

8 Sittig DF, Singh H. A new sociotechnical model for studying health information technology in complex adaptive healthcare systems. *Qual Saf Health Care* 2010;19(Suppl 3):i68–74.

9 Dyrbye LN, Shanafelt TD, Sinsky CA, *et al*. Burnout among health care professionals: a call to explore and address this underrecognized threat to safe, high-quality care. *NAM Perspectives* 2017;7.

10 Shanafelt TD, Dyrbye LN, West CP. Addressing physician burnout: the way forward. *JAMA* 2017;317:901–2.

11 Westbrook JI, Coiera E, Dunsmuir WTM, *et al*. The impact of interruptions on clinical task completion. *Qual Saf Health Care* 2010;19:284–9.

12 Weigl M, Müller A, Vincent C, *et al*. The association of workflow interruptions and hospital doctors' workload: a prospective observational study. *BMJ Qual Saf* 2012;21:399–407.

13 Weigl M, Beck J, Wehler M, *et al*. Workflow interruptions and stress at work: a mixed-methods study among physicians and nurses of a multidisciplinary emergency department. *BMJ Open* 2017;7:e019074.

14 Ancker JS, Edwards A, Nosal S, *et al*. Effects of workload, work complexity, and repeated alerts on alert fatigue in a clinical decision support system. *BMC Med Inform Decis Mak* 2017;17.

15 Bell NR, Dickinson JA, Grad R, *et al*. Understanding and communicating risk: measures of outcome and the magnitude of benefits and harms. *Can Fam Physician* 2018;64:181–5.

16 Chernev A, Böckenholt U, Goodman J. Choice overload: a conceptual review and meta-analysis. *J Consum Psychol* 2015;25:333–58.

17 Bollen D, Knijnenburg BP, Willemsen MC, *et al*. Understanding choice overload in recommender systems. In: *Proceedings*

of the fourth ACM conference on recommender systems, 2010: 63–70.

18 Caban JJ, Gotz D. Visual analytics in healthcare—opportunities and research challenges. *J Am Med Inform Assoc* 2015;22:260–2.

19 Yu K-H, Kohane IS. Framing the challenges of artificial intelligence in medicine. *BMJ Qual Saf* 2019;28:238–41.

20 Challen R, Denny J, Pitt M, *et al*. Artificial intelligence, bias and clinical safety. *BMJ Qual Saf* 2019;28:231–7.

21 Wyatt JC, Altman DG. Commentary: prognostic models: clinically useful or quickly forgotten? *BMJ* 1995;311:1539–41.

22 Holzinger A, Biemann C, Pattichis CS, *et al*. What do we need to build explainable AI systems for the medical domain? *arXiv preprint* 2017.

23 Beam AL, Kohane IS. Big data and machine learning in health care. *JAMA* 2018;319:1317–8.

24 Topol E. *Deep medicine: how artificial intelligence can make healthcare human again*. New York: Basic Books, 2019.