



Digital Oncology

Ethical challenges of machine learning and deep learning algorithms

Almost every journal or technology news publication we read today mentions the next artificial intelligence-based algorithm that can make diagnoses or predict outcomes. Various academic and privately funded organisations are currently working on machine learning or deep learning applications that will either help physicians' decision making or work independently, and possibly even replace physicians in some settings. In this context, there is a growing need for the next generation of physicians to understand the basis of how such applications are developed. Indeed, to be considered truly ethical, an algorithm or system development and implementation should consider the views of all parties involved—physicians, patients, parents (in the case of children) or carers (in the case of older people), academic researchers, commercial entities working on developing new solutions, and even insurance providers. This is indeed a high bar, but one that should be established as a standard for industry to follow. Also, there is currently no defined framework to help decide who is responsible when a machine learning algorithm or a robotic system designed to use these tools gets things wrong. Is the end user (ie, the physician or the hospital) always responsible, or does the failure get classified as a product liability resulting from malfunctioning equipment?

There is often a lack of clear information about the failure rate of tools currently approved for use in clinical practice and the specific groups of patients or data points where they fail. As an example, the US Food and Drug Administration (FDA) approved a tool for automated detection of wrist fractures in 2018. This tool detects distal radius fractures in adult wrists and is an adjunct tool not intended to replace a clinician's review of the radiograph. Although this might be a useful application in settings where there is a shortage of radiologists, detection of radius fractures is not considered among the top challenges faced by radiologists or orthopaedists experienced in reading wrist radiographs. The more pertinent question is how the tool performs in the subset of patients with osteoporotic bones that human readers find challenging. There is also a danger of over-dependence on artificial intelligence tools, overlooking pertinent clinical findings that were not made available or not assessed by the algorithm. To mitigate this risk, it would be prudent to ensure that a trained physician checks the result generated by the algorithm. As an example, a tool recently approved by the FDA for the detection of diabetic retinopathy requires further

evaluation by a trained eye-care provider. Indeed, training of physicians making this final arbitration in the future should include details of how to vet these tools, so that physicians can override or accept an outlier result, and train and improve the algorithm. The ability of a machine to learn when a case is truly normal depends in many cases on the training data sets used, and the current state of knowledge. As with humans, advances in medical knowledge might mean that algorithms will need updating at regular intervals and these algorithms could become outdated as soon as they are validated. Such adaptive algorithms that become better over time are the future, but there are potential issues. For instance, users might have to pay regularly for updates. In addition to the increased costs, there is a danger that data from at least a proportion of patients will be analysed with an outdated algorithm, if their facility is not able to install or deploy the new version for some reason.

As patient information is converted into data points, there is a danger of dehumanisation of data. To avoid this, there should be close collaboration between researchers, clinicians, patients, and patient advocacy groups on these topics. As patient information gathered during the doctor-patient interaction is converted into data for use by machine learning algorithms, there is a danger of dilution of the doctor-patient relationship. The patient

Dr Sanjay P Prabhu
Department of Radiology,
Boston Children's Hospital and
Harvard Medical School, Boston,
MA, 02115, USA
sanjay.prabhu@childrens.harvard.edu

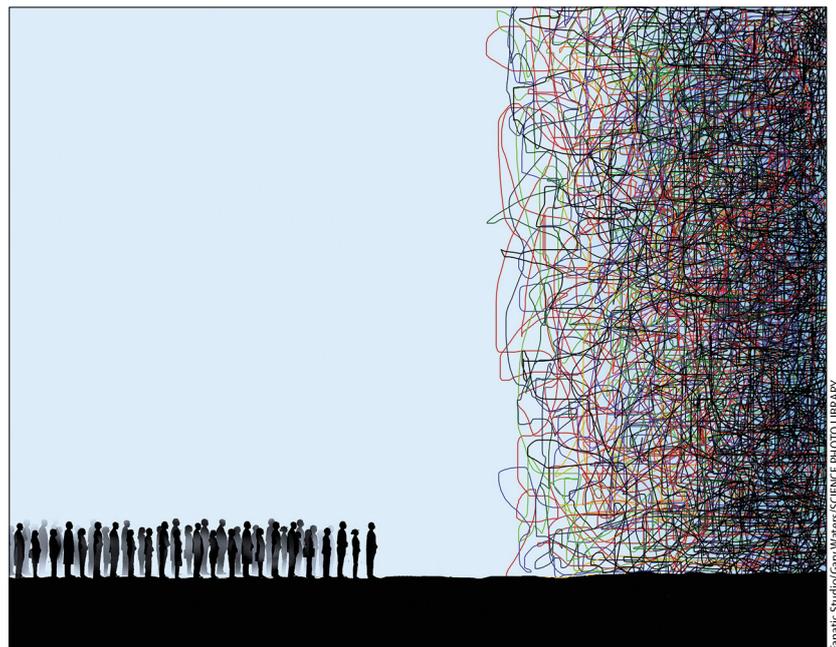
I declare no competing interests

For more on the **promise of machine learning applications to identify features in images not apparent to humans** see
J Digit Imaging 2017; **30**: 391

For more on **patients' consent** see
J Am Coll Radiol 2018; **15**: 580–86

For more on **automated tools for detection of diabetic retinopathy** see
npj Digital Medicine 2018; **1**: 40

For more on **applications for automated detection of focal lesions causing seizures** see
Neuroimage Clin 2016; **14**: 18–27



may well ask if they are providing data to the physician or to a machine via the doctor. Patients should be informed about how their information will be used, and informed consent should be obtained to use their information for the development of future applications. Currently, some institutions obtain consent from patients when they enter the facility to collect and store blood samples and other health-care data as part of the development of biobanks. These biobanks data can generally be used for any type of research on human health, including genomics, biomarker analyses, epidemiology and outcomes research, and the creation of cell lines. In other cases, large technology companies are gaining access to patient data by working with health-care facilities, buying access to image storage repositories or working with insurance companies that have obtained patients' consent for data use. Although there are opt-out clauses in most of these consent forms, it is not always clear to patients and providers what they are consenting to when they sign the informed consent forms.

Patients who do not consent to their data being used for the development of artificial intelligence algorithms or object to the use of their data for this purpose should continue to be accommodated in the health-care system. Opting out should not mean losing out. They should also be allowed to revoke access to their data if they feel that these are being used in ways that were not made expressly clear when initial consent was obtained.

There are no clear regulations about how the data collected are used and shared with third parties, and we currently rely on Institutional Review Boards (IRBs) and ethics committees in institutions to help guide the process. The checks and balances of IRBs are not available in all hospitals and all parts of the world. It is possible that we might require laws such as General Data Protection Regulation (GDPR) tailored for the use of data of individuals in health-care applications to ensure that companies and individuals working in this area remain cognisant of patient privacy at all times. Patient advocacy groups and medical associations will have to work with and guide law makers to make these laws fair to both patients and researchers trying to develop new tools.

As more solutions are developed to assist or replace human readers, one could argue that it is unethical not to allow access to machine learning tools to patients and providers in low-resource settings if the cost of a suite of solutions is lower than employing human readers and if these tools can help shorten diagnostic times and ensure prompt management. On a personal note, in my own practice, my colleagues and I use semi-automated co-registration and subtraction tools to compare the size of ventricles or tumour masses at different timepoints in our patients, as we find this is more accurate than manual visual assessment. I cannot see why all patients should not have access to these tools if they have been shown to be better than human assessment alone.

As the performance of artificial intelligence tools starts surpassing those of humans performing a similar task, patients will request direct or indirect access to such applications. For example, if a patient with a tumour wants to know the likelihood of malignancy, they may wish to use a suite of applications made available online before consenting to biopsy. As another example, patients may request access to online applications for automated detection of a focal abnormality causing seizures, where there is a large variation in the accuracy of imaging studies when scans are read by radiologists with varying levels of experience.

Ethical issues surrounding machine learning and artificial intelligence are complex and deserve a lot of thought before incorporating them into routine clinical care. It would be prudent to include patient advocates and technically savvy ethics experts in decision making at all levels to pre-emptively come up with solutions to potential ethical challenges in this field. And on a practical note, I suggest checking in with your own institution and confirming that they have a clear policy of how they use and share patient data, and educating your patients about the data use consent forms they are asked to sign.

Sanjay P Prabhu



Digital Oncology

Crowdfunding for cancer research: the TRACe campaign as an example

Biosciences Laboratory (GG, PF), and Research and technology transfer office (PM, LB), Istituto Scientifico Romagnolo per lo Studio e la Cura dei Tumori (IRST) IRCCS, 47014, Meldola,

Over the past 7 years, research funding has witnessed a changing trend. Besides available competitive research funding such as Horizon 2020 (Europe), National Institutes of Health (USA), National Health Service (UK) grants, and other public research funding programmes

around the world, including those from charitable organisations, researchers have started to engage in crowdfunding activities to look for alternative sources of funding. Management of a scientific grant, based on developing a technical project proposal that will be