



Since January 2020 Elsevier has created a COVID-19 resource centre with free information in English and Mandarin on the novel coronavirus COVID-19. The COVID-19 resource centre is hosted on Elsevier Connect, the company's public news and information website.

Elsevier hereby grants permission to make all its COVID-19-related research that is available on the COVID-19 resource centre - including this research content - immediately available in PubMed Central and other publicly funded repositories, such as the WHO COVID database with rights for unrestricted research re-use and analyses in any form or by any means with acknowledgement of the original source. These permissions are granted for free by Elsevier for as long as the COVID-19 resource centre remains active.

## Machine learning for COVID-19—asking the right questions



Published Online  
July 10, 2020  
[https://doi.org/10.1016/S2589-7500\(20\)30162-X](https://doi.org/10.1016/S2589-7500(20)30162-X)

COVID-19, caused by the novel severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2), has put health-care systems worldwide into crisis.<sup>1</sup> The speed with which health-care resources have been consumed has in some countries exceeded supply of personal protective equipment and ventilators, the unprecedented need for the latter as a result of life-threatening respiratory failure that characterises severe disease.

Among the principal diagnostic imaging modalities, both chest x-ray and CT have quickly produced a large amount of data on COVID-19, enabling the development of machine learning algorithms, a form of artificial intelligence (AI). Well before the COVID-19 pandemic, enthusiasm around machine learning-based technology in medical imaging had notably increased. Now, huge datasets emerging from China, and increasingly from European countries, have generated numerous publications reporting AI applications in COVID-19.

What remains to be seen is how many of these applications will prove to be clinically useful. The first step to achieving this goal is to define the clinical need for which a solution will improve or transform clinical care. The danger is that without expert clinical oversight, applied AI research might result in solutions' looking for problems: a form of supply trying to find demand, rather than the other way around. Although AI-based medical imaging research is published frequently, the number of systems validated in clinical trials and implemented in clinical practice is comparatively small.<sup>2</sup> Google Deepmind's collaboration with Moorfields Eye Hospital (London, UK)<sup>3</sup> might be considered a prototypical example of expert clinical oversight driving technical innovation, where a high accuracy, automated solution to analyse optical coherence tomography retinal scans was developed to address the huge volume of scans done globally each year. Similar, problem-focused applications of machine learning are now being implemented in the National Health Service, UK, including Microsoft's InnerEye technology for radiotherapy planning to save time and Heart Flow's machine-learning tool for 3D coronary modelling from cardiac CT, which provides decision support to clinicians assessing a patient's need for coronary angiography.<sup>4</sup>

However, for COVID-19, research questions risk focusing too much on generating novel machine learning

models without fully considering its practical application and potential biases.

Occasionally the speed and accuracy of machine learning algorithms are reported on the basis of performance in clinical scenarios that do not accurately reflect clinical practice. Sometimes comparisons between algorithm and human performance are unbalanced. In most cases a computer has been trained to detect a specific abnormality (eg, COVID-19-related parenchymal disease), whereas a radiologist is usually responsible for detecting any abnormality (including incidental findings such as pulmonary nodules or pulmonary emboli).

Machine learning-based CT analysis has also been suggested as a promising screening tool for COVID-19, and in at least one study outperformed viral real-time PCR testing.<sup>5</sup> However, these results need to be interpreted cautiously. Studies done during a pandemic are inherently hampered by artificially high disease prevalence and the selected nature of participants, whose disease severity warranted hospital admission and CT evaluation. Ideally, algorithms need to be trained on the full spectrum of disease, including asymptomatic and early-stage cases, if CT interpretation by machine learning can be applied to real-world data with confidence. Furthermore, a consensus must be reached on what the best data labelling strategy might be: are only patients with positive real-time PCR considered to be infected with SARS-CoV-2? Should data labelling incorporate multidisciplinary information such as the presence of a cough or fever? How does a study participant's exposure to an infected relative or household member alter algorithm training? In most cases, machine learning algorithms will be developed on retrospective, clinically indicated data that are often imperfect. However, rather than invalidate model training, incorporating all the statistical noise associated with real-world clinical data in model training might improve an algorithm's clinical applicability.

Undoubtedly, COVID-19 offers many exciting opportunities for applied AI research. But research questions must be prioritised according to their probable clinical effect. As we learn more about the natural history of COVID-19, it has become apparent that the disease progresses in stages.<sup>6</sup> The need to

pre-empt deterioration and personalise preventative interventions have emerged as a priority. Currently, imaging research has focused on diagnosis on the basis of appearances once the disease has progressed. Detection of disease at the earliest stages, when initiation of appropriate therapy is likely to be most effective, would be more useful. CT also has a well established role as a prognostic tool in many diffuse lung diseases, particularly when combined with clinical data. This finding is of importance in COVID-19; given that a primary concern for health-care providers is becoming overwhelmed by patients requiring intensive care and ventilatory support, accurate prognostication is arguably a more pressing clinical problem than diagnosis.<sup>7</sup> For COVID-19, training an algorithm to predict outcomes such as mortality, intensive care unit admission, or need for mechanical ventilation could have considerable clinical effect.<sup>8-10</sup>

An untapped resource in patients with COVID-19 is the availability of chest x-rays at multiple time points early in a patient's disease. In other respiratory disorders, short-term disease behaviour is the strongest predictor of long-term outcome.<sup>9</sup> By incorporating sequential chest x-rays into model training, novel imaging features of progressive disease, including features inaccessible to the human eye, might be disclosed. More generally, although patients with comorbid conditions represent a population at high clinical risk, it is currently not possible to identify patients with no underlying health issues but who are also likely to develop progressive disease. The availability of objective stratification tools to rapidly assess a patient would assist frontline health-care workers in making difficult decisions about the allocation of scarce resources.<sup>8</sup>

If the history of pandemics is any guide, moments of crises can accelerate innovation, in part by creating permissive environments for collaboration. The COVID-19 pandemic is no different, as shown by the rapid setup of responses, such as the Imaging COVID-19 AI initiative, a multi-centre European project for pooling CT images across vast and diverse populations to power machine learning research. However, the starting point for this project has again focused largely on diagnosis; reorienting research questions towards pressing clinical problems including outcome prediction, with use of baseline or short-term data, might be more fruitful.

Machine learning algorithms are often modular, meaning that new algorithms generated during this pandemic might be successfully repurposed for other pulmonary diseases in the future.

Lastly, balancing light-touch regulation—as has increasingly been the position advocated by governments—with robust ethical standards is essential to build an environment that enables rapid review and appropriate ethical approval. However, we must avoid rushing through unproven solutions in response to the COVID-19 pandemic. As always, there will be a balance of risk and rapidity, and the key to optimising this balance is to define the needs for which solutions will have the greatest clinical value. With the right collaboration between clinical and machine learning expertise, the current public health crisis might mark the beginning of a decade when AI in health care delivers on its promises of wide, transformative clinical impact.

PB declares no competing interests. NSP reports personal fees from Google and Synergix, outside of the submitted work. SLFW reports personal fees from Sanofi-Aventis, Roche, Open Source Imaging Consortium, Bracco, and FLUIDDA; grants and personal fees from Boehringer Ingelheim and Galapagos; and grants from National Institute for Health Research, outside the submitted work. We would like to thank the National Institute of Health Research Imperial Biomedical Research Centre.

Copyright © 2020 The Author(s). Published by Elsevier Ltd. This is an Open Access article under the CC BY 4.0 license.

**Patrik Bachtiger, Nicholas S Peters, \*Simon LF Walsh**  
s.walsh@imperial.ac.uk

National Heart and Lung Institute, Imperial College London, London SW3 6LY, UK

- 1 Tanne JH, Hayasaki E, Zastrow M, Pulla P, Smith P, Garcia Rada A. Covid-19: how doctors and healthcare systems are tackling coronavirus worldwide. *BMJ* 2020; **368**: m1090.
- 2 Wiens J, Saria S, Sendak M, et al. Do no harm: a roadmap for responsible machine learning for health care. *Nat Med* 2019; **25**: 1337–40.
- 3 De Fauw J, Ledsam JR, Romera-Paredes B, et al. Clinically applicable deep learning for diagnosis and referral in retinal disease. *Nat Med* 2018; **24**: 1342–50.
- 4 Douglas PS, De Bruyne B, Pontone G, et al. 1-year outcomes of FFRCT-guided care in patients with suspected coronary disease: the PLATFORM study. *J Am Coll Cardiol*; **68**: 435–45.
- 5 Li L, Qin L, Xu Z, et al. Artificial intelligence distinguishes COVID-19 from community acquired pneumonia on chest CT. *Radiology* 2020; published online March 19. DOI:10.1148/radiol.20200905.
- 6 Weiss P, Murdoch DR. Clinical course and mortality risk of severe COVID-19. *Lancet* 2020; **395**: 1014–15.
- 7 Phua J, Weng L, Ling L, et al. Intensive care management of coronavirus disease 2019 (COVID-19): challenges and recommendations. *Lancet Respir Med* 2020; **8**: 506–17.
- 8 Colombi D, Bodini FC, Petrini M, et al. Well-aerated lung on admitting chest CT to predict adverse outcome in COVID-19 pneumonia. *Radiology* 2020; published online April 17. DOI:10.1148/radiol.202001433.
- 9 Walsh SLF, Wells AU, Sverzellati N, et al. An integrated clinicoradiological staging system for pulmonary sarcoidosis: a case-cohort study. *Lancet Respir Med* 2014; **2**: 123–30.
- 10 Causey JL, Zhang J, Ma S, et al. Highly accurate model for prediction of lung nodule malignancy with CT scans. *Sci Rep* 2018; **8**: 1–12.