# The challenges of deploying artificial intelligence models in a rapidly evolving pandemic

The attention and resources of AI researchers have been captured by COVID-19. However, successful adoption of AI models in the fight against the pandemic is facing various challenges, including moving clinical needs as the epidemic progresses and the necessity to translate models to local healthcare situations.

Yipeng Hu, Joseph Jacob, Geoffrey J. M. Parker, David J. Hawkes, John R. Hurst and Danail Stoyanov

The COVID-19 pandemic, caused by the severe acute respiratory syndrome coronavirus 2, emerged into a world that was seeing rapid developments in artificial intelligence (AI) based on big data, computational power and neural networks. In recent years, the gaze of AI researchers has increasingly turned to applications in healthcare. Inevitably, there has been much interest in exploring the potential for AI to support the response to the pandemic across a wide range of clinical and societal challenges[1], for instance in disease forecasting, disease surveillance and antiviral drug discovery. However, to date AI has had surprisingly little impact on the management of COVID-19. This Comment focuses on examining the possible reasons behind the lack of successful adoption of AI models specifically for COVID-19 diagnosis and prognosis in frontline healthcare services. We highlight the moving clinical needs that models have had to address at different stages of the epidemic, and explain the importance of translating models to reflect local healthcare environments. We argue that both basic and applied research are essential to accelerate the potential of AI models, and this is particularly so during a rapidly evolving pandemic. This perspective on the response to COVID-19 may provide a glimpse into how the global scientific community should react to combat future disease outbreaks more effectively.

## The evolving clinical need

The clinical management of COVID-19 has spanned various stages including anticipation, early detection, containment and mitigation, together aiming towards eventual eradication[2]. Each stage differs in its measured and actual disease prevalence, which directly impacts the availability of clinical resources, and over a matter of weeks, clinical priorities can fluctuate rapidly. Priorities may range from providing robust diagnoses, to maintaining infection control and ensuring availability of
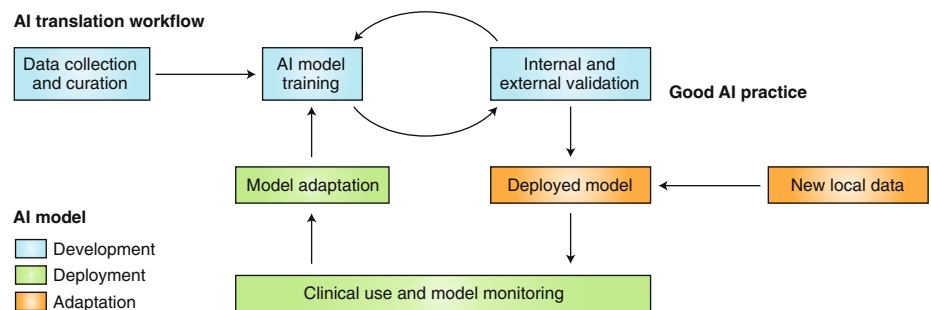


**Fig. 1 | Translating AI models.** Illustration of a typical AI translation workflow, including AI model development, deployment and adaptation (or model update).

facilities for mechanical ventilation. These rapid changes, occurring in tandem with enhanced knowledge of virus behaviour and increasing availability of supporting data, have meant that the outputs required of predictive AI models need to constantly evolve. Accordingly, the AI models that are most urgently needed and can feasibly be built are likely to be different at each epidemic stage.

**The anticipation and early detection stages.** With a relatively low number of positive cases and many potentially asymptomatic cases during the early stages of the pandemic, a highly sensitive diagnostic AI model to detect COVID-19 would have been useful. The lack of pre-existing data from this new disease means the feasibility of building such new AI models to determine diagnosis or prognosis is a challenge, one that could be addressed by an AI community focused on breaking the existing barriers between data domains using machine learning. Generalizing AI models to unseen data (inference), data coming from different distributions (domain adaptation, transfer learning) and data with limited or no labels (semi- or unsupervised learning)[3] are all priority areas in the technical development

of AI. The early stages of a new disease have been described as overlooked periods in the general management of infectious diseases[4]. The AI community should design strategies and methodologies for rapid deployment in the event of future epidemic threats, to make data collection, model training, testing and wide deployment as efficient as possible next time.

**The containment and mitigation stages.** During the containment and mitigation of COVID-19, data have become increasingly available with the exponential growth of confirmed positive cases. During this time it is essential to rapidly curate sizable training datasets and develop stable, well-performing AI models that can respond to emerging clinically urgent needs, such as rapid, consistent patient triage at scale across a health service.

Reverse transcription polymerase chain reaction (RT-PCR) tests via nasopharyngeal swabbing have nearly 100% specificity and are considered the diagnostic ground-truth for COVID-19. However, RT-PCR has limited negative predictive value with variable availability and diagnostic speed. Alternative methodologies for diagnosing COVID-19 include medical imaging techniques such as computed tomography

(CT) and chest radiographs[5]. Some groups have also explored point-of-care ultrasound, albeit with limitations[6]. Driven by data availability, the focus of AI work in COVID-19 has centred on RT-PCR-labelled diagnostic models[7], or the automated evaluation of clinical/imaging features — for example, lung involvement on CT imaging[8]. Those developing AI-assisted diagnostic tools must recognize that very high diagnostic accuracies are required to demonstrate added value above and beyond existing clinical imaging and RT-PCR tests.

It is also important to question which prognostic outcomes require greatest prioritization during this period. The majority of existing AI models aim to predict hospitalization and mortality[9] using predictors such as age, gender, blood biomarkers, pre-existing co-morbidities and imaging[5]. In resource-constrained clinical environments, there is great value in predicting resource consumption as a 'surrogate' prognostic outcome, as a lack of personal protective equipment, for example, can directly affect community prognoses[10]. Intuitive candidate outcome measures for AI models might include time spent on mechanical ventilators or within intensive care units. But as knowledge of COVID-19 has grown, early intubation of a patient has diminished in priority in the care pathway. Similarly, with limitations in resources, pragmatic choices have had to be made regarding patient selection for intensive care unit admission. Evolving management strategies such as these have a real-time impact on the outcomes that AI models aim to predict. Disease progression (or regression following treatment) models can be trained using time series data, such as longitudinal CT images[11], to quantify the likelihood of developing severe pneumonia and acute myocardial injury, two leading causes of mortality[12], and the cytokine release syndrome. A lesson from the COVID-19 pandemic has been that AI models motivated merely by the practical convenience of acquiring available labelled data has had limited clinical value.

**The eradication stage and beyond.** At the later eradication stage, constraints in data availability, development time and clinical resources would gradually be eased. The number of positive cases can drop quickly. Yet the need for a real-time, convenient, highly sensitive screening tool may persist to control transmission and judiciously recognize potential outbreaks.

The requirements for prognostic tools may shift to the identification of patients at risk of developing long-term health problems such as pulmonary fibrosis.

Indeed cardiopulmonary, neurological[13] and urological[14] damage are all being recognized following COVID-19 infection. Given the potentially significant health service resource requirements that may result from long-term complications across large swathes of the population, the post-acute phase of COVID-19 will be a critical clinical research area, where AI models may play a central role.

### Translating AI models
A typical AI translation workflow (Fig. 1) includes model development, model deployment and model adaptation (or model update). The COVID-19 AI research efforts have been concentrated primarily on new model development and the urgency brought about by the pandemic must not override the stringent requirements for clinical deployment[15]. Despite time pressures, rigorous validation is key to ensuring that safety and efficacy are tested; models must be validated before initial deployment and continuously monitored and adapted when implemented in local healthcare environments and as outcome likelihoods change due to evolving patient management strategies. Failing to adhere to such practice will impede translation and compromise the impact of AI on clinical needs.

**Pre-deployment validation.** Recent COVID-19 AI models have been criticized for a lack of transparency in development and a high likelihood of bias towards non-representative patient populations[5]. Limitations in data availability and quality can be the inherent cause of problems — for example, validation datasets with unrealistically high numbers of control cases acquired at the start of an outbreak or extremely low numbers of control cases at the peak of the outbreak. These models are unlikely to be directly useful in all stages of the pandemic due to potential bias.

Best practices in rigorous design and analysis of experiments should be adopted for AI model validation. In addition, model interpretation methods help to explain the reasoning of the predictions[16,17], and may also indicate when certain data-driven methods are unlikely to generalize[18]. Model transparency could also be key to addressing regulatory and ethical issues[19,20].

**Local adaptation.** It is not uncommon to find that an AI model trained with data from one healthcare centre, or even from multiple centres, does not generalize as well at a new centre. For example, the accuracy of chest X-ray detection, represented by the area under the receiver operating

characteristic curve, was significantly reduced from 0.93 in a multi-centre internal validation to 0.82 on external validation data[18]. The practice of pre-deployment external validation reduces the risk of this overfitting problem based on the assumption that external data represent new local data. However, for each individual healthcare environment, local data are likely to have unique characteristics due to centre-specific acquisition features, equipment and protocols, all of which may have differing clinical constraints and requirements[21]. Moreover, temporal differences in data may increase, adversely affecting model accuracy, as the demographic and immunity landscape and clinical practice shift between different stages of the pandemic[22]. AI models therefore should have a continuous monitoring and adaptation strategy to these changing data to maintain their predictive accuracy.

Most proposed AI approaches for COVID-19 diagnosis/prognosis have so far been 'locked' algorithms that do not facilitate future adaptation. Model-adapting methods from other medical applications should be tested and integrated in these developments, such as transfer learning[23] and model retraining with a small local dataset. Recently, the US Food and Drug Administration has proposed a new approach to allow AI-based software to adapt and improve from real-world use[24], paving the regulatory pathway to address these local adaptation needs.

### Conclusion
The COVID-19 pandemic has presented numerous challenges to virtually every section of society in all geographic locations. AI can be an enabling technology to support urgent clinical needs in disease diagnosis and prognosis but is reliant on appropriate infrastructure, data management and translational pathways. New international cross-disciplinary collaborations, carefully identifying time-, course- and region-dependent clinical actions in response to COVID-19 can benefit from scientifically sound AI model development, validation and deployment to support local healthcare providers. Safe and responsible translation is the only way to realize the promise of AI models to contribute to combating the current coronavirus pandemic, its aftermath and potential future clustered outbreaks or comparable healthcare emergencies. ❒

Yipeng Hu [ID][1,2,3] ✉, Joseph Jacob[1,4], Geoffrey J. M. Parker [ID][1,5,6], David J. Hawkes[1,2,3], John R. Hurst[4] and Danail Stoyanov[1,2,5]

[1]UCL Centre for Medical Image Computing, University College London, London, UK. [2]Wellcome/EPSRC Centre for Interventional and Surgical Sciences, University College London, London, UK. [3]Department of Medical Physics and Biomedical Engineering, University College London, London, UK. [4]UCL Respiratory, University College London, London, UK. [5]Department of Computer Science, University College London, London, UK. [6]Bioxydyn Limited, Manchester, UK.
✉e-mail: yipeng.hu@ucl.ac.uk

## References

1. Bullock, J., Pham, K. H., Lam, C. S. N. & Luengo-Oroz, M. Preprint at https://arxiv.org/abs/2003.11336 (2020).
2. *Managing Epidemics: Key Facts About Major Deadly Diseases* (World Health Organization, 2018).
3. Cheplygina, V., de Bruijne, M. & Pluim, J. P. W. *Med. Image Anal.* **54**, 280–296 (2019).
4. Webby, R. J. & Webster, R. G. *Science* **302**, 1519–1522 (2003).
5. Wynants, L. et al. *BMJ* **369**, m1328 (2020).
6. Soldati, G. et al. *J. Ultrasound Med.* https://doi.org/10.1002/jum.15285 (2020).
7. Li, L. et al. *Radiology* https://doi.org/10.1148/radiol.2020200905 (2020).
8. Huang, L. et al. *Radiol. Cardiothoracic Imaging* **2**, e200075 (2020).
9. Yuan, M., Yin, W., Tao, Z., Tan, W. & Hu, Y. *PLoS ONE* **15**, e0230548 (2020).
10. Rubin, G. D. et al. *Radiology* https://doi.org/10.1148/radiol.2020201365 (2020).
11. Pan, F. et al. *Radiology* https://doi.org/10.1148/radiol.2020200370 (2020).
12. Zheng, Y.-Y., Ma, Y.-T., Zhang, J.-Y. & Xie, X. *Nat. Rev. Cardiol.* **17**, 259–260 (2020).
13. Filatov, A., Sharma, P., Hindi, F. & Espinosa, P. S. *Cureus* **12**, e7352 (2020).
14. Li, Z. et al. Preprint at https://doi.org/10.1101/2020.02.08.20021212 (2020).
15. Zagury-Orly, I. & Schwartzstein, R. M. *New Engl. J. Med.* https://doi.org/10.1056/NEJMp2009405 (2020).
16. Lundberg, S. M. & Lee, S. I. *In Advances in Neural Information Processing Systems* **30**, 4765–4774 (2017).
17. Kim, B. et al. In *Proc. 35th Int. Conf. Machine Learning* 2668–2677 (PMLR, 2018).
18. Zech, J. R. et al. *PLoS Med.* **15**, https://doi.org/10.1371/journal.pmed.1002683 (2018).
19. Ribeiro, M. T., Singh, S. & Guestrin, C. In *Proc. 22nd ACM SIGKDD Int. Conf. Knowledge Discovery and Data Mining* 1135–1144 (ACM, 2016).
20. Goodman, B. & Flaxman, S. *AI Mag.* **38**, 50–57 (2017).
21. Beede, E. et al. in *Proc. 2020 CHI Conf. Human Factors in Computing Systems* 1–12 (ACM, 2020).
22. Chen, M., Hao, Y., Hwang, K., Wang, L. & Wang, L. *IEEE Access* **5**, 8869–8879 (2017).
23. Van Opbroek, A., Ikram, M. A., Vernooij, M. W. & De Bruijne, M. *IEEE Trans. Med. Imaging* **34**, 1018–1030 (2014).
24. *Proposed Regulatory Framework for Modifications to Artificial Intelligence/Machine Learning (AI/ML)-Based Software as a Medical Device (SaMD)* (US Food and Drug Administration, 2019).