

# Integrating artificial intelligence in bedside care for covid-19 and future pandemics

**Michael Yu and colleagues** examine the challenges in developing AI tools for use at point of care

**T**he covid-19 pandemic created unprecedented challenges for both clinicians and healthcare institutions. Adapting to a rapidly emerging disease while facing staff and material shortages prompted difficult decisions on how best to allocate resources. Artificial intelligence (AI) rapidly moved to the forefront of the effort to adapt our healthcare systems to coping with covid-19. Hundreds of new models were developed, promising best solutions for all aspects of patient care from diagnostics to therapeutics and logistics. Yet only a small minority of these models were deployed, and none became widely adopted.<sup>1,2</sup> We argue that the covid-19 pandemic exposed flaws in the technological, institutional, and ethical foundations upon which AI must build to considerably improve bedside care. If AI is to be part of a rapid response to future health crises, the challenges that it faced during the covid-19 pandemic must be carefully analysed and overcome.

AI is a branch of computer science that uses data and algorithms to extract meaning in a way that is characteristic

of intelligent beings—that is, turning data into effective decision making processes. Research applications of AI in medicine have already emerged far and wide—for example, in drug discovery and modelling of complex biological systems. By contrast, efforts to integrate AI into everyday clinical care have had minimal success, despite the comparatively simple nature of the problems: optimising patient trajectories, maximising use of existing facilities, or determining when and how to reallocate resources. We surmise that this translational gap, which was magnified by the covid-19 pandemic, is due to the nature of the underlying data, the infrastructure through which they emerge, and the human context in which they occur. By understanding the influence of these factors on the chances of success of AI, healthcare systems can improve their readiness for future crises.

## Development challenges

Many high quality data are essential for training an AI model, during which patterns are identified from observations, and are developed into a generalisable model for prediction. The covid-19 pandemic disclosed significant differences in the ability of institutions to provide data rapidly in standardised formats, suitable for training AI models. Even within a single institution, anomalous events, such as the start of a pandemic, can limit the availability of relevant data. These problems are compounded when collaboration extends beyond institutional, provincial, and national borders—often forcing researchers to choose between large datasets of superficial information, and smaller datasets that delve into more detail. Failure to procure data of sufficient quality or quantity can make the development of an AI model challenging, and lead to biases and erroneous conclusions.

Bedside clinical care relies on inputs from many inter-related processes, which often straddle disjointed data systems. These data systems must be reconciled before they can be used for training AI models. Data might originate from patient monitoring devices, observations entered by workers into electronic health records,

or known constraints on human and material resources. Manual entry might be unavoidable to capture data that are known only to patients, care givers, or hospital decision makers. Existing systems might need to be adapted to emergency practices, such as the repurposing of units to create temporary intensive care facilities. Regardless of the initial source of data, interoperability is critical. Institutions should promote efforts to adopt well established standard terminologies, in order to aid data quality assessments and collaboration. Data standardisation is a cumbersome and rate limiting process, which should be started well before downstream applications.

If a particular event under study is rare, cooperation between institutions might be essential to make the use of AI models feasible. For example, early in the pandemic, the modest number of patients with covid-19 treated at individual institutions was insufficient to train robust models for prediction of patient trajectories and outcomes.<sup>3,4</sup> Pooling patient level data from many institutions to create larger datasets—the traditional approach in clinical research—proved to be impractical, owing to time consuming ethical and legislative considerations. Here, AI offers several interesting new solutions. “Few-shot learning” (making predictions based on a limited number of samples) is an approach that uses readily available data (eg, historical information) to make generalisations about unseen data—for example, allowing the development of a covid-19 diagnostic model using a small number of training cases.<sup>5</sup> Federated learning is another approach that relies on sharing models, rather than data, across institutions. This strategy allows pooling of knowledge without exchange of sensitive information.<sup>6</sup> Much like data standardisation, these approaches require a high level of technological preparedness to be ready for use during a health crisis.

Training with a sufficient quantity of high quality data does not guarantee that AI models will generalise well. Similar concepts might be defined differently across institutions: for instance, operations

## KEY MESSAGES

- Despite substantial research, artificial intelligence (AI) has had a limited effect on bedside care during the covid-19 pandemic
- To develop AI tools for bedside care and hospital management, researchers require standards for integrating data from disparate sources, and software that can process data in real time
- Variations in healthcare institution policies, resources, expertise, patient populations and differences between prepandemic and postpandemic operations require careful consideration in generalising AI solutions
- Engaging AI researchers, healthcare providers, ethicists, and other experts is necessary to develop standards for establishing when it is appropriate to deploy AI tools and to integrate AI into existing workflows

categorised as “urgent” might imply different types of care across different hospitals. Recent work has attempted to create models that account for such differences.<sup>7</sup> These challenges are magnified across geographical areas, where concepts might need to be translated between languages, and adjustments made for baseline patient characteristics and technological capabilities. In some instances, differences in technological capabilities could prevent low income countries from benefiting from AI, or require specialised approaches towards knowledge transfer. Federated learning might improve model generalisability, because it allows for localisation when part of a model is trained locally. The performance trade-offs of institution specific approaches should be weighed against more generic AI solutions.<sup>8</sup>

Undoubtedly, the overarching solution to the challenges described in this section is to enhance coordination and knowledge sharing among researchers working on AI models for bedside care. A living systematic review of AI models pertaining to covid-19, as of August 2021, identified 27 models for disease progression.<sup>9</sup> Another systematic review, published in March 2021, evaluated 62 studies involving the use of chest computed tomography for prediction of covid-19 disease progression.<sup>10</sup> Both reviews found methodological flaws and risks of bias in almost all studies, attributing this, at least in part, to lack of coordination. To meet systemic demands, and to better guide the direction of AI research, institutions should support efforts to move collaboration beyond recognising common features and weakness across studies towards a deeper synthesis of best practices and lessons learnt to actively improve future research. The covid-19 pandemic has bolstered AI workshops, including notable efforts in the United Kingdom.<sup>2</sup> These workshops present a key opportunity to develop solid guidance for AI solution developers, and to create a stronger, more tightly knit community.

#### Requirements for deployment and adoption

To achieve a sustainable impact, researchers into AI should look beyond model development and consider how solutions can be practically and ethically implemented at the bedside. This approach demands a broader perspective that ensures integration with hospital systems, satisfies ethical standards to safeguard patients, and adapts to existing workflows in a way that acknowledges and leverages clinical

expertise. If AI researchers do not adapt their work to real world clinical contexts, they risk producing models that are irrelevant, infeasible, or irresponsible to implement. One challenge is upgrading hospitals systems to support trained AI models. For example, diagnostic models for early detection of lung disease might need to be integrated with picture archiving and dictation systems. Upgrading these systems can be difficult or unwise during a health crisis, suggesting that health institutions would benefit from investing in these infrastructure upgrades before the demand arises.

AI also needs to be guided by clear ethical standards.<sup>11 12</sup> For example, researchers recently discovered unintentional racial biases built into decision making algorithms used to recommend care in the United States.<sup>13</sup> Pressure to develop AI models quickly heightens the risks of exacerbating inequalities within and across borders.<sup>12</sup> Efforts to establish standards for research involving AI have progressed, including the CONSORT-AI extension for clinical trials using AI,<sup>14</sup> and developing work on TRIPOD-AI for reporting guidelines.<sup>15</sup> These standards will enhance transparency and reproducibility, which had remained wanting up to this point.<sup>16</sup> As with diagnostic tests, medical devices, and therapeutics, translation of AI from bench to bedside should follow evidence based clinical evaluation standards. Faced with the pressures of a pandemic, these standards are crucial to ensuring more responsible model development.

To be used at the bedside, AI models must also achieve clinical acceptance through mutual, effective knowledge sharing between AI researchers and clinicians. As the potential impact of AI at the bedside increases, medical educators should consider introducing the basics of data science and machine learning to all clinicians.<sup>17</sup> Meanwhile, AI researchers can develop a better understanding of clinical needs and avoid developing black box technologies as much as possible. Thus adaptation is required, such as developing more interpretable models, following the approach of explainable AI.<sup>18-20</sup> The acceptance and usability of AI models can also be improved by working with clinicians to design models that supplement—rather than replace—clinical judgment.<sup>21</sup> Additionally, AI researchers should integrate other tools in clinical decision making, when appropriate. For example, predictions from an AI model can be enhanced when embedded in a decision support system that strengthens “what if”

scenarios, allowing the user to explore alternative strategies to treat patients when beds in the intensive care unit are at capacity.<sup>22</sup> Disruption to clinical processes caused by the covid-19 pandemic should be seen as an opportunity for clinicians and health institutions to rethink their practices.

Patient acceptance of AI at the bedside must be earned through all stages of development and implementation. This requires guiding AI research to take account of patient values, and deal with legitimate concerns about a technology that can magnify biases and inequalities in the healthcare system.<sup>12</sup> A critical first step towards this goal is to participate in cross sectional and intersectional research to understand the diversity of patient perspectives across factors such as sex, age, wealth, race and ethnicity, and health status. Subsequently, guidelines will need to incorporate such perspectives into enforceable policies that can help to govern the design, development, implementation, maintenance, and oversight of AI as applied to bedside care.

#### Enabling robust, agile, responsible AI at the bedside

AI has yet to produce a single widely accepted tool for bedside care of patients with covid-19, reflecting gaps in research, technical details, and the policy environment surrounding it rather than a failure in the field. A considerable amount of work remains to be done before solid foundations are laid for deployment of AI at the bedside.

Researchers in AI can benefit from a more expansive view of the problems they are solving, by considering the work of other researchers, and the full decision making and technical context in which their solutions are to be deployed. Developing networks with other AI researchers, clinicians, hospital staff, and patients can guide research efforts towards high impact areas. During a pandemic, establishment of these networks can be used to accelerate research, while acting as a safeguard against potential oversights.

On the technical side, health institutions need to take advantage of crisis free periods to develop the infrastructure required—including efforts to promote standardisation and interoperability. Adopting, promoting, and enhancing standards such as the Fast Healthcare Interoperability Resources, which is used to capture institutional data, helps to make AI more readily transferable, and supports

multisite research.<sup>23</sup> Coordination across institutions is also fundamental, and many initiatives have been spurred on by the covid-19 pandemic, such as OpenSafely in the UK or CODA-19 in Canada.<sup>24 25</sup>

Finally, clear guidelines are required for AI research and deployment. To achieve clear guidelines requires a large scale effort to understand the diverse values and needs of those who might be affected by AI at the bedside, including patients, care givers, and hospital administrators. This process should be followed by the creation of enforceable standards, involving multistage evaluations of AI models, which can serve as guiding principles for model development and deployment.

Creating this research, technical, and policy infrastructure requires a substantial effort, although one that is already under way. We are confident that AI researchers already recognise the need to move the focus beyond individual projects, and towards developing translational tools that can guide AI research into clinical practice.

We thank the Consortium Santé Numérique de l'Université de Montréal, especially Kamran Afzali, Pascale Beliveau, Khedidja Seridi, and Yves Terrat, for their insights, review, and comments on this article.

**Contributors and sources:** MY, MC, and LM conceptualised and wrote the first draft of the manuscript. SW and JR contributed to the reflection on clinical experience and evidence based policy making and helped to contextualise this work with the WHO research agenda related to artificial intelligence. All authors provided scientific, methodological, and content expertise. All authors reviewed, edited, and approved the final version of the manuscript. MC and LM are guarantors. SW and JR are staff members of WHO. The authors alone are responsible for the views expressed in this article and they do not necessarily represent the decisions, policy, or views of WHO.

**Competing interests:** We have read and understood BMJ policy on declaration of interests and have no relevant interests to declare.

**Provenance and peer review:** Commissioned; externally peer reviewed.

This collection of articles was proposed by WHO Department of Digital Health and Innovation and commissioned by *The BMJ*. *The BMJ* retained full editorial control over external peer review, editing, and publication of these articles. Open access fees were funded by WHO.

Michael Yu, methodologist<sup>1</sup>

An Tang, professor of radiology<sup>2</sup>

Kip Brown, research associate<sup>1</sup>

Rima Bouchakri, data scientist<sup>1</sup>

Pascal St-Onge, health data science programme manager<sup>1</sup>

Sheng Wu, technical officer<sup>3</sup>

John Reeder, director, special programme for research and training in tropical diseases<sup>3</sup>

Louis Mullie, intensivist<sup>1</sup>

Michaël Chassé, deputy scientific director of health data science<sup>1</sup>

<sup>1</sup>Center for Integration and Analysis of Medical Data (CITADEL), Department of Medicine, University of Montreal, Centre de Recherche du Centre Hospitalier de l'Université de Montréal, Montreal, Canada

<sup>2</sup>Department of Radiology, Centre Hospitalier de l'Université de Montréal (CHUM), Canada

<sup>3</sup>World Health Organization, Geneva, Switzerland

Correspondence to: M Chassé

michael.chasse.med@ssss.gouv.qc.ca



## OPEN ACCESS

This is an Open Access article distributed under the terms of the Creative Commons Attribution IGO License (<https://creativecommons.org/licenses/by-nc/3.0/igo/>), which permits use, distribution, and reproduction for non-commercial purposes in any medium, provided the original work is properly cited.



- Nicholson W. Artificial intelligence in the medical system: four roles for potential transformation. *Public Law Leg Theory Res Pap Ser* 2019;631:122-32.
- Alan Turing Institute. In: von, Brozyskowski I, Mazumder A, Mateen B, et al. eds. *Data science and AI in the age of COVID-19*. 2021. [https://www.turing.ac.uk/sites/default/files/2021-06/data-science-and-ai-in-the-age-of-covid\\_full-report\\_2.pdf](https://www.turing.ac.uk/sites/default/files/2021-06/data-science-and-ai-in-the-age-of-covid_full-report_2.pdf)
- Boserup B, McKenney M, Elbuli A. The impact of the COVID-19 pandemic on emergency department visits and patient safety in the United States. *Am J Emerg Med* 2020;38:1732-6. doi:10.1016/j.ajem.2020.06.007
- Rosenbaum L. The untold toll—the pandemic's effects on patients without covid-19. *N Engl J Med* 2020;382:2368-71.
- Sawyer D, Fiaidhi J, Mohammed S. Few shot learning of covid-19 classification based on sequential and pretrained models: a thick data approach. In: 2021 IEEE 45th annual computers, software, and applications conference (COMPSAC). IEEE, 2021:1832-6. <https://ieeexplore.ieee.org/document/9529642/>
- Kaissis G, Ziller A, Passerat-Palmbach J, et al. End-to-end privacy preserving deep learning on multi-institutional medical imaging. *Nat Mach Intell* 2021;3:473-84. doi:10.1038/s42256-021-00337-8.
- Khan A, Atzori M, Otálora S, Andrearczyk V, Müller H. Generalizing convolution neural networks on stain color heterogeneous data for computational pathology. In: Tomaszewski JE, Ward AD, eds. *Medical imaging 2020: digital pathology*. SPIE, 2020:26.
- Futoma J, Simons M, Panch T, Doshi-Velez F, Celi LA. The myth of generalisability in clinical research and machine learning in health care. *Lancet Digit Health* 2020;2:e489-92.
- Wynants L, Van Calster B, Collins GS, et al. Prediction models for diagnosis and prognosis of covid-19: systematic review and critical appraisal. *BMJ* 2020;369:m1328. doi:10.1136/bmj.m1328
- Roberts M, Driggs D, Thorpe M, et al. Common pitfalls and recommendations for using machine learning to detect and prognosticate for COVID-19 using chest radiographs and CT scans. *Nat Mach Intell* 2021;3:199-217. doi:10.1038/s42256-021-00307-0.
- Cave S, Whittlestone J, Nyrop R, O hEigeartaigh S, Calvo RA. Using AI ethically to tackle covid-19. *BMJ* 2021;372:n364. doi:10.1136/bmj.n364
- Leslie D, Mazumder A, Peppin A, Wolters MK, Hagerty A. Does "AI" stand for augmenting inequality in the era of covid-19 healthcare? *BMJ* 2021;372:n304. doi:10.1136/bmj.n304
- Obermeyer Z, Powers B, Vogeli C, Mullainathan S. Dissecting racial bias in an algorithm used to manage the health of populations. *Science* 2019;366:447-53.
- Ibrahim H, Liu X, Rivera SC, et al. Reporting guidelines for clinical trials of artificial intelligence interventions: the SPIRIT-AI and CONSORT-AI guidelines. *Trials* 2021;22:11. doi:10.1186/s13063-020-04951-6
- Collins GS, Dhiman P, Andaur Navarro CL, et al. Protocol for development of a reporting guideline (TRIPOD-AI) and risk of bias tool (PROBAST-AI) for diagnostic and prognostic prediction model studies based on artificial intelligence. *BMJ Open* 2021;11:e048008. doi:10.1136/bmjopen-2020-048008
- Haibe-Kains B, Adam GA, Hosny A, et al. Massive Analysis Quality Control (MAQC) Society Board of Directors. Transparency and reproducibility in artificial intelligence. *Nature* 2020;586:E14-6. doi:10.1038/s41586-020-2766-y
- Meskó B, Görög M. A short guide for medical professionals in the era of artificial intelligence. *NPJ Digit Med* 2020;3:126. doi:10.1038/s41746-020-00333-z.
- Markus AF, Kors JA, Rijnbeek PR. The role of explainability in creating trustworthy artificial intelligence for health care: a comprehensive survey of the terminology, design choices, and evaluation strategies. *J Biomed Inform* 2021;113:103655. doi:10.1016/j.jbi.2020.103655
- Chassagnon G, Vakalopoulou M, Battistella E, et al. AI-driven quantification, staging and outcome prediction of COVID-19 pneumonia. *Med Image Anal* 2021;67:101860. doi:10.1016/j.media.2020.101860
- Drouin A, Letarte G, Raymond F, Marchand M, Corbeil J, Lavolette F. Interpretable genotype-to-phenotype classifiers with performance guarantees. *Sci Rep* 2019;9:4071. doi:10.1038/s41598-019-40561-2
- Robbins R. 'Human experts will make the call': Stanford launches an accelerated test of AI to help care for Covid-19 patients. *STAT* 2020. <https://www.statnews.com/2020/04/01/stanford-artificial-intelligence-coronavirus/>
- Sperrin M, Diaz-Ordaz K, Pajouheshnia R. Invited commentary: treatment drop-in-making the case for causal prediction. *Am J Epidemiol* 2021;190:2015-8. doi:10.1093/aje/kwab030
- Health Level Seven International. HL7 FHIR. <http://www.hl7.org>
- The Data Lab. OpenSAFELY. <https://www.opensafely.org/>
- CODA-19: Collaborative data analysis to improve clinical care in patients with covid-19. 2021. <https://www.coda19.com/>

Cite this as: *BMJ* 2021;375:e068197

<http://dx.doi.org/10.1136/bmj-2021-068197>