



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.



Digital medicine Is my cough COVID-19?

For more on **Digital medicine** see **Comment Lancet** 2016; **388**: 740 and **Perspectives Lancet** 2020; **396**: 1479

Imagine that even if you have no symptoms of COVID-19, the sound of your forced cough transmitted to your smartphone or smart speaker, processed by an algorithm, could provide a 98.5% accurate diagnosis. That's what a study involving more than 4000 people suggested might be possible. And it could be done anytime, free of charge, with immediate turnaround of results. It's one of many proposed uses of artificial intelligence (AI) for COVID-19. However, such technology will clearly require further research and independent replication to be refined, accepted, or implemented.

Indeed, replication is a major concern in the use of AI in medicine, as exemplified by two recent studies. Earlier this year, a study of mammograms from more than 25 000 women in the UK and over 3000 women in the USA suggested an AI algorithm increased the accuracy of breast cancer diagnosis compared with radiologists. But other researchers questioned these findings, asserting that a lack of transparency—in sharing of the code and sufficiently documented methods—make the results irreproducible. Similarly, an algorithm-based mathematical modelling approach for predicting COVID-19 mortality using three biomarkers in 485 patients suggested 90% accuracy. Multiple research teams subsequently tested this model and found the accuracy of mortality prediction was poor. The lack of replication, or what can be considered external validation here, was not due to insufficient transparency, but rather, like many other such studies based on a very small cohort, an inadequately supported conclusion.

A better attempt with AI for predicting COVID-19 deterioration was a model that used data from thousands of patients in almost 600 hospitals in three provinces in China with nearly 90% accuracy. The problem in extrapolating that finding, and its replications, is related to the crucial issue that the outputs of deep neural networks are fully dependent on the inputs. We have seen time and time again that race, ethnicity, geography, location, and other demographic factors influence an algorithm's performance. Any AI model can only be deemed to apply to those patients whose data were used as its basis.

That nuance can be missed, as has been seen with studies that have used AI for interpreting chest CT scans in patients with COVID-19. Accurately distinguishing coronavirus from other causes for pneumonia or using the scan to make the diagnosis instead of a virus test has been advocated. Some of these studies are large enough with test and validation cohorts that are robust. However, the studies have been done in places where COVID-19 was highly prevalent. Replication in regions with a low prevalence of COVID-19 has not been attempted.

Bypassing the need for both replication and proof that an AI model works before implementation is another problem. For instance, an automated predictive model for clinical deterioration in hospital (not COVID-19 related) among more than 325 000 patients at 19 hospitals was implemented, and compared with the period before its use; the findings indicated a lower hospital mortality and intensive care unit admission rate, along with a shorter length of stay, with the model. Yet without a randomised trial, it is hard to assess the veracity of these results.

The same concerns apply when patients with COVID-19 are triaged to home instead of being admitted to the hospital on the basis of an AI algorithm. Many health systems in the USA are using algorithms for patients with a wearable sensor that captures continuous oximetry, body temperature, heart rate and rhythm, respirations, and mobility. Such remote monitoring of patients with mild to moderate COVID-19 has potential, but this approach has been implemented without a single peer-reviewed publication or preprint, without even an attempt at replication. There are limited prospective studies and even fewer randomised trials of AI in medicine, emphasising our need for much stronger and concerted efforts to develop the robust evidence for clinical use.

So back to the cough and COVID-19 concept. A pioneer of deep neural network AI, Geoffrey Hinton, has said "Deep learning is going to be able to do everything." And there have been media headlines about AI's role in the COVID-19 response. The hype for AI was profound even before the pandemic; it has been magnified since. Until we have definitive evidence and replication and external validation, with all the caveats discussed here, that AI can be used to provide an accurate diagnosis of COVID-19 from a forced cough, we should resist the notion—no matter how alluring it seems.

Eric J Topol

Scripps Research Translational Institute and the Scripps Research Institute, La Jolla, CA 92037, USA
etopol@scripps.edu

Further reading

- Laguarta J, Huetto F, Subirana B. Covid-19 artificial intelligence diagnosis using only cough recordings. *IEEE Open J Eng Med Biol*; published online Sept 29. <https://doi.org/10.1109/OJEMB.2020.3026928>
- McKinney SM, Karthikesalingam A, Tse D, et al. Reply to: transparency and reproducibility in artificial intelligence. *Nature* 2020; **586**: e17-18
- Topol EJ. Welcoming new guidelines for AI clinical research. *Nat Medicine* 2020; **26**: 1318-20
- Yan L, Zhang HT, Goncalves J, et al. An interpretable mortality prediction model for COVID-19 patients. *Nat Mach Intell* 2020; **2**: 283-88
- 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
- EJT is supported by the US National Institutes of Health/ National Center for Advancing Translational Sciences grant UL1TR001114



Faith-Haig/Scripps Research Translational Institute