Medical Xpress
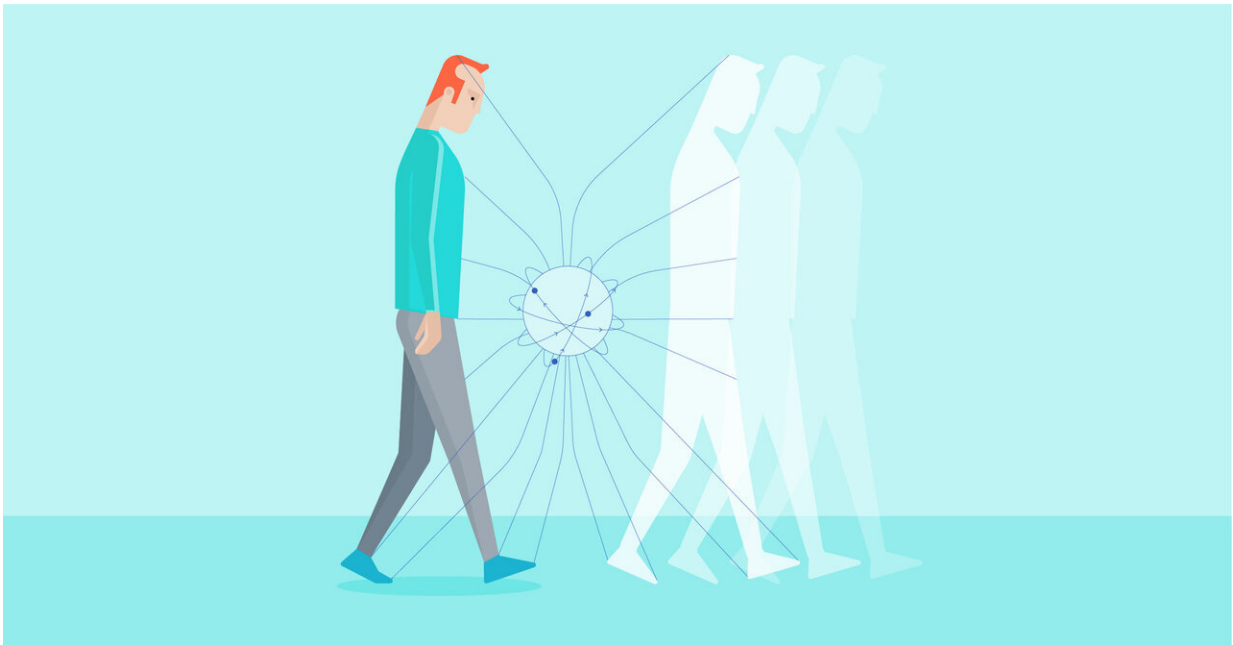
AI with 'imagination' could help doctors with diagnosis, particularly for complex case

August 11 2020

Babylon's new Artificial Intelligence with 'imagination' could help doctors with diagnosis. Credit: Babylon Health

Doctors are overworked and in short supply around the globe, but they could soon be assisted by machine learning to reduce errors in primary care. AI symptom checkers are tremendously valuable in providing medical information and safe triaging advice to users. However, none of them performs diagnoses like a doctor. Unlike doctors, existing symptom checkers provide advice based on correlations alone—and

correlation is not causation. Researchers at Babylon have, for the first time that we know of, used the principles of causal reasoning to enable AI to diagnose written test cases.

The researchers used a new approach, known as causal machine learning—which is gaining increased traction in the AI community—to act as an "imagination" so the AI could consider what symptoms it might see if the patient had an illness different to the one it was considering. The peer-reviewed research, published in *Nature Communications*, shows that disentangling correlation from causation makes the AI significantly more accurate.

Dr. Jonathan Richens, Babylon scientist and lead author, said, "We took an AI with a powerful algorithm and gave it the ability to imagine alternate realities and consider if a symptom would be present if it was a different disease. This allows the AI to tease apart the potential causes of a patient's illness and score more highly than over 70% of the doctors on these written test cases."

Dr. Ali Parsa, CEO and founder of Babylon, said, "Half the world has almost no access to healthcare. We need to do better. So it's exciting to see these promising results in test cases. This should not be sensationalized as machines replacing doctors, because what is truly encouraging here is for us to finally get tools that allow us to increase the reach and productivity of our existing healthcare systems. AI will be an important tool to help us all end the injustice in the uneven distribution of healthcare, and to make it more accessible and affordable for every person on Earth."

A pool of over 20 Babylon GPs created 1,671 realistic written medical cases—these included typical and atypical examples of symptoms for more than 350 illnesses. Each case was authored by a single doctor and then verified by multiple other doctors to ensure it represented a realistic

diagnostic case. A separate group of 44 Babylon GPs were then each given at least 50 written cases (the mean was 159) to assess. The doctors listed the illnesses they considered most likely (on average returning 2.58 potential diseases for each diagnosis). They were measured for accuracy by the proportion of cases where they included the true disease in their diagnosis. Babylon's AI took the same tests and used both an older algorithm based on correlations created specifically for this research, and the newer, causal one. For each test, the AI could only report as many answers as the doctor had.

The doctors had a mean score of 71.40% (± 3.01%) and ranged from 50-90%. The older correlative algorithm performed on par with the average doctor, achieving 72.52% (± 2.97%). The new causal algorithm scored 77.26% (± 2.79%) which was higher than 32 of the doctors, equal to 1, and lower than 11.

Dr. Tejal Patel, associate medical director and GP, Babylon, said, "I'm excited that one day soon this AI could help support me and other doctors reduce misdiagnosis, free up our time and help us focus on the patients who need care the most. I look forward to when this type of tool is standard, helping us enhance what we do."

Dr. Saurabh Johri, chief scientist and author, Babylon, added, "Interestingly, we found that the AI and doctors complemented each other—the AI scored more highly than the doctors on the harder cases, and vice versa. Also, the algorithm performed particularly well for rare diseases which are more commonly misdiagnosed, and more often serious. Switching from using correlations improved accuracy for around 30% of both rare and very-rare conditions."

It is not necessary to alter the underlying models of disease that an AI uses in order to get an improvement in accuracy. It is a benefit that would apply to existing correlative algorithms, including those outside of

the medical setting.

Dr. Ciaran Lee, study author, formerly of Babylon and honorary lecturer at UCL, said, "Causal machine learning allows us to ask richer, more natural questions about medicine. This method has huge potential to improve every other current symptom checker, but it can also be applied to many other problems in healthcare and beyond—that's why causal AI is so impressive, it's universal."

This technology paves the way for a future partnership between clinicians and AI that will speed up a doctor's diagnosis, improve accuracy, free up time for clinicians and improve patient outcomes and patient experiences. It has the potential to augment the work of clinicians and continue to drive a better healthcare system for patients.

This new causal algorithm is not yet present in Babylon's publicly available app. It will only be released after further development and testing, and once it has met all necessary regulatory approvals in the UK and other markets where it will be released.

More information: Improving the accuracy of medical diagnosis with causal machine learning, *Nature Communications* (2020). DOI: 10.1038/s41467-020-17419-7

Provided by Babylon