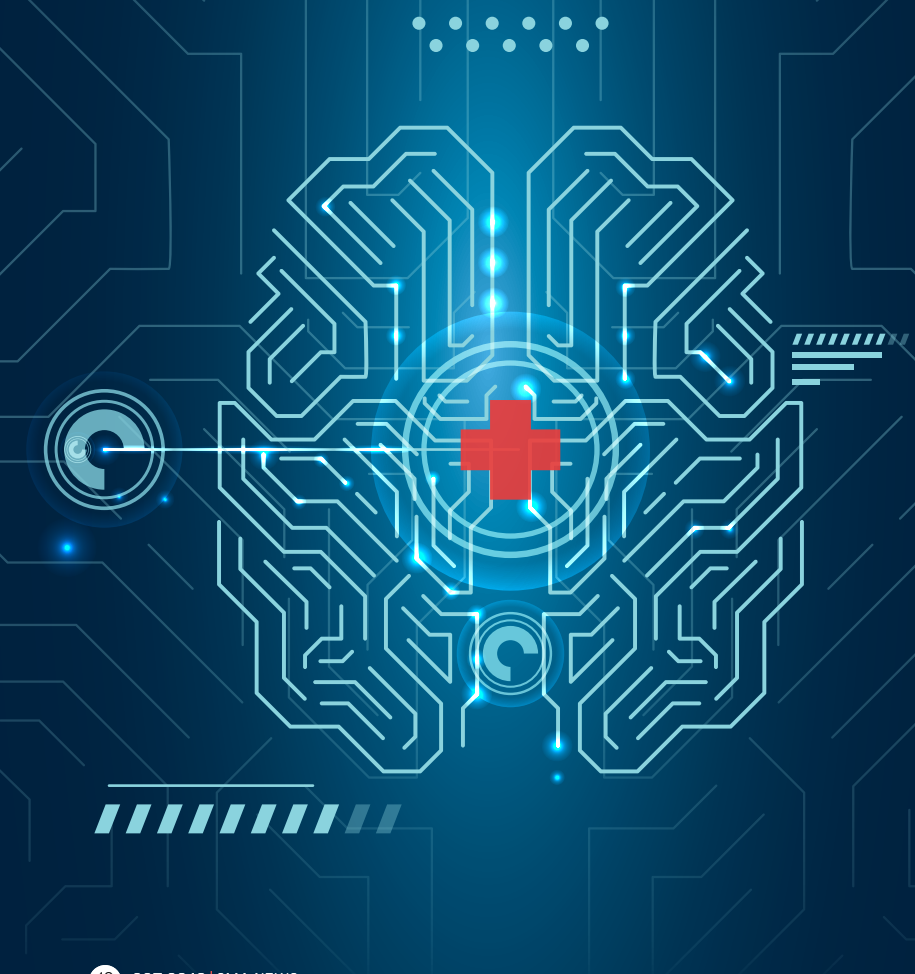
Are We Entering a Super-Physician Age?

Text by Prof Henk Schmidt



Stop training radiologists. That was the controversial claim made by British cognitive psychologist and computer scientist Geoffrey Hinton in 2016. In 2017, a paper published in *Nature* showed that artificial intelligence (AI) can learn to classify skin cancer as accurately as dermatologists,[1] and earlier this year, this claim was advanced by a further study that found that AI outperformed experts in the diagnosis of melanomas.[2]

Do these advances herald the substitution of human physicians with AI? Or can we harness its computing powers to improve overall patient care through a new hybrid approach?

The application of AI in medicine has so far concentrated most heavily on the area of visual diagnoses because images contain both the explicit and implicit cues needed to arrive at an accurate diagnosis. With deep-learning capabilities, AI programmes can potentially excel, if fed with enough and varied data representing all disease states and permutations.

The potential of AI in non-visual diagnostics

But what about instances where the information needed to make an accurate diagnosis relies on verbal communication of a patient's experience of a disease? In a recent study in the UK,[3] patients diagnosed with coronary artery disease were asked to describe their chest pain. While more than half described a "tight" or "tightening" sensation, a range of other phrases was also used, including "pressure", "constriction", "sharp", "stabbing" and more vague terms including "dull", "stitch-y", and "tingling".

While these terms, though varied, fit the angina canon,[a] and could, therefore, be used as source information for an AI application, other patients used more abstract terminology, including similes, synonyms and slang. Some patients even described symptoms not suggestive of angina, but that they felt were important. Yet in each case, the diagnosis was the same. In Singapore, we have an additional complexity layer: multiple cultural backgrounds

and languages that affect how these experiences are lived and described.

The reality is that diseases do not always present in a standard way. There is enormous variety both in terms of the stage and presentation of a disease, as well as a patient's experience of it. Each consult demands that a physician applies his or her clinical reasoning skills anew to deduce the right diagnosis. For that, experiential knowledge is crucial as book knowledge alone is insufficient.

A physician becomes an expert through experience built up over years and decades of seeing many patients and the same diseases in many forms. The experienced physician learns to incorporate knowledge that he or she may not even be aware of. For example, the way a patient smells can be an indicator of certain diseases, and the way they behave or move as they walk into the consultation room can yield clues about their health. How this tacit knowledge can be incorporated is a challenge that AI has yet to tackle, and some experts in the field, such as the late American philosopher Hubert Dreyfus, argue that such knowledge cannot be effectively captured through formal rules.

A boost to experienced physicians

If it cannot supplant the physician, what role is there for AI? AI has the potential to act as an effective support for physicians, particularly in situations where experienced physicians cannot arrive at a diagnosis using their intuition (or System 1 thinking). It has long been accepted in medicine that being able to think fast (ie, using System 1) is the hallmark of expertise. If an experienced radiologist sees an X-ray and knows very quickly what is going on, he or she is most likely right.

But when a diagnosis is not evident to the experienced radiologist, he or she has to engage System 2 thinking – a more considered and step-by-step approach to analysing the information presented. This type of thinking has been shown to result in a greater number of diagnostic errors. In such a case, a support system that has seen thousands of different slides representing the many different representations of disease may help by offering differential diagnoses for further analysis by the radiologist.

It could, therefore, be argued that an effective AI support system could boost experienced physicians' ability to make accurate diagnoses by suggesting differential diagnoses that the physician may have overlooked.

Ultimately, this could result in overall safer patient care.

The wait for the super physician

Such diagnostic support programmes already exist. The user types in the signs and symptoms observed, and the computer produces a number of alternative hypotheses. But these are cumbersome and error-prone, and many residents complain bitterly about such systems' unwieldiness.

Unlike domains which behave according to a set of rules that are clearly defined, like mathematics and aviation for example, medicine operates in a grey space with poorly defined parameters that are, in addition, highly subjective and variable depending on who is affected by what disease, how that affects the individual and how they, in turn, describe it. This continues to be the main hurdle facing AI in diagnostic medicine.

Even in aviation, the autopilot has not replaced the pilot. Every cadet has to complete a set number of flight hours before he or she is licensed to captain a plane. The same holds true for physicians. Medical students need to understand the relationship of the signs and symptoms with the underlying pathophysiological processes. They need to accumulate experience to develop expertise, and it is then, and only then, that AI diagnostic support devices can make a significant difference.

I do believe that there will come a time when such support systems will have developed sufficiently to become useful, but we still have some way to go. In the 1980s, when the first rudimentary AI programmes first emerged, everyone was certain that AI would be better than humans within a few years.

However, it took another 20 years before AI beat a chess grandmaster, and chess is played according to fixed rules and moves. Of course, AI has changed gear since then. Now, systems can learn. If you give them enough pictures, they become better and more accurate. But for diagnostic support systems to reach this level of accuracy, we really need another 50 years. ◆

References

1. Esteva A, Kuprel B, Novoa RA, et al. Dermatologist-level classification of skin cancer with deep neural networks. Nature 2017; 542:115-8.

2. Haenssle HA, Fink C, Schneiderbauer R, et al. Man against machine: diagnostic performance of a deep learning convolutional neural network for dermoscopic melanoma recognition in comparison to 58 dermatologists. Ann Oncol 2018; 29(8):1836-42.

3. Jones MM, Somerville C, Feder G, Foster G. Patients' descriptions of angina symptoms: a qualitative study of primary care patients. Br J Gen Pract 2010; 60:735-41.

Note

a. The angina canon refers to the reliance of interpreting the well-described set of symptoms of cardiac and non-cardiac chest pain of patients.

Prof Schmidt is professor of psychology at Erasmus University, Netherlands, and visiting professor of Medical Education at the Lee Kong Chian School of Medicine (LKCMedicine). He will present about "How students learn" at the inaugural Transform MedEd conference, a joint conference by LKCMedicine and Imperial College School of Medicine, on 9 to 10 November 2018. To register for the conference, visit https://www.transformmeded.org/.