

Software is Eating the Clinic

When I was a medical student close to graduation, Sam Shuster then Professor of Dermatology in Newcastle, drew my attention to a paper that had just been published in *Nature*. The paper, from the laboratory of Robert Weinberg, described how DNA from human cancers could transform cells in culture (1). I tried reading the paper, but made little headway because the experimental methods were alien to me. Sam did better, because he could distinguish the underlying melody from the supporting orchestration. He told me that whilst there were often good papers in *Nature*, perhaps only once every 10 years or so would you read a paper that would change both a field *and* the professional careers of many scientists. He was right. The paper by Weinberg was one of perhaps fewer than a dozen that defined an approach to the biology of human cancer that still resonate 40 years later.

Revolutionary papers in science have one of two characteristics. They are either conceptual, offering a theory that is generative of future discovery – think DNA, and Watson and Crick. Or they are methodological, allowing what was once impossible to become almost trivial – think DNA sequencing or CRISPR technology. Revolutions in medicine are slightly different, however. Yes, of course, scientific advance changes medical practice, but to fully understand clinical medicine we need to add a third category of revolution. This third category comes from papers that change the everyday lives of what doctors do and how they work. Examples would include fibre-optic instrumentation and modern imaging technology. To date, dermatology has escaped such revolutions, but a paper recently published in *Nature* suggests that our time may have come (2).

The core clinical skill of the dermatologist is categorising morphological states in a way that informs prognosis with, or without, a therapeutic intervention. Dermatologists are rightly proud of these perceptual skills, although we have little insight as to how this expertise is encoded in the human brain. Nor should we be smug about our abilities as, although the domains are different, the ability to classify objects in the natural world is shared by many animals, and often appears effortless. Formal systems of education may be human specific, but the cortical machinery that allows such learning, is widespread in nature.

There have been two broad approaches to try and imitate these skills *in silico*. Either particular properties (shape, colour, texture etc.) are first explicitly identified and, much as we might add variables in a linear regression equation, the information used to try and discriminate between lesions in an explicit way. Think of the many papers using rule-based strategies such as the ABCD system (3). This is obviously *not* the way the human brain works: a moment's reflection about how fast

an expert can diagnose skin cancers and how limited we are in being able to handle formal mathematics, tells us that human perceptual skills do not work like this.

There is an alternative approach, one to some extent that almost seems like magic. The underlying metaphor is as follows. When a young child learns to distinguish between cats and dogs, we know the language of explicit rules is not used: children cannot handle multidimensional mathematical space or complicated symbolic logic. But feedback, in terms of what the child thinks, allows the child to build up his or her own model of the two categories (cats versus dogs). With time, and with positive and negative feedback, the accuracy of the perceptual skills increase – but without any formal rules that the child could write down or share. And of course, since it is a human being we are talking about, we know all of this process takes place within and between neurons.

Computing scientists started to model the way that they believed collections of neurons worked over 4 decades ago. In particular, it became clear that groups of *in silico* neurons could order the world based on positive and negative feedback. The magic is that we do not have to explicitly program their behaviour, rather they just learn, but – since this is not magic after all – we have got much better at building such self-learning machines. (I am skipping any detailed explanation of such 'deep learning' strategies, here). What gives this field its current immediacy is a combination of increases in computing power, previously unimaginable large data sets (for training), advances in how to encode such 'deep learning', and wide potential applicability – from email spam filtering, terrorist identification, online recommendation systems, to self-driving cars. And medical imaging along the way.

In the *Nature* paper by Thrun and colleagues (2) such 'deep learning' approaches were used to train computers based on over 100,000 medical images of skin cancer or mimics of skin cancer. The inputs were therefore 'pixels' and the diagnostic category (only). If this last sentence does not shock you, you are either an expert in machine learning, or you are not paying attention. The 'machine' was then tested on a new sample of images and – since modesty is not a characteristic of a young science – the performance of the 'machine' compared with over 20 board certified dermatologists. If we use standard receiver operating curves (ROC) to assess performance the machine equalled if not out-performed the humans.

There are of course some caveats. The dermatologists were only looking at single photographic images, not the patients (4); the images are possibly not representative of the real world; and some of us would like to know more about the exact comparisons used. However, I would argue that there are also many reasons for imagining that the paper may underestimate the power of this approach:

it is striking that the machine was learning from images that were relatively unstandardised and perhaps noisy in many ways. And if 100,000 seems large, it is still only a fraction of the digital images that are acquired daily in clinical practice.

It is no surprise that the authors mention the possibilities of their approach when coupled with the most ubiquitous computing device on this planet – the mobile phone. Thinking about the impact this will have on dermatology and dermatologists would require a different sort of paper from the present one but, as Marc Andreessen once said (4), ‘software is eating the world’. Dermatology will survive, but dermatologists may be on the menu.

Jonathan REES, Prof, FMedSci

Grant Chair of Dermatology, Edinburgh Medical School: Clinical Sciences, University of Edinburgh, Dermatology, Rm 4.018. Lauriston Building, Lauriston Place, Edinburgh, EH3 9HA, UK. E-mail: reestheskin@me.com

REFERENCES

1. Shih C, Padhy LC, Murray M, Weinberg RA. Transforming genes of carcinomas and neuroblastomas introduced into mouse fibroblasts. *Nature* 1981; 290: 261–264.
2. Esteva A, Kuprel B, Novoa RA, Ko J, Swetter SM, Blau HM, Thrun S. Dermatologist-level classification of skin cancer with deep neural networks. *Nature* 2017; 542: 115–118.
3. Aldridge RB, Zanotto M, Ballerini L, Fisher RB, Rees JL. Novice identification of melanoma: not quite as straightforward as the ABCDs. *Acta Derm Venereol* 2011; 91: 125–130.
4. Gaudy-Marqueste C, Wazaefi Y, Bruneu Y, Triller R, Thomas L, Pellacani G, Grob JJ. Ugly duckling sign as a major factor of efficiency in melanoma detection. *JAMA Dermatol* 2017 Feb 8. [Epub ahead of print].
5. Marc Andreessen, *Wall Street Journal*, August 20th, 2011. Available from: <https://www.wsj.com/articles/SB10001424053111903480904576512250915629460>.



Are the dermatologists on the menu?