The first machine learning application in healthcare was approved by the FDA as recently as 2019 to analyse MRI images of the heart. In oncology, applications that make use of artificial intelligence (AI) are appearing thick and fast, from improving diagnostics to developing new biomarkers. The algorithms being developed could lead to significant advances and cost savings, but this is a field with a history of over-promising. So what can AI do in both diagnostics and therapeutics and how is it currently helping decision making? And what role will AI play in future developments of precision oncology? Could it provide the tool to interpret the massive amounts of genomic data becoming available and provide us with insights that can ultimately improve patient care?

Artificial intelligence has a rocky history dating back to the 1950s. In 1997, IBM's Deep Blue computer defeated chess champion Garry Kasparov and by 2011, IBM's new Watson supercomputer was able to win the US$1m prize in the US game-show Jeopardy. This quiz competition presented contestants with general knowledge clues in the form of answers, and they had to phrase their responses in the form of a question. This progress came from advances in machine learning, where a computer can be 'trained' to find patterns on sets of data and then apply this knowledge to new data. This has now advanced to deep learning, whereby systems are able to improve their performance or 'learn' when exposed to sets of data and can essentially programme themselves.

AI in diagnostics and treatment

One of the first areas in oncology to take advantage of AI was diagnostic radiology, where it has the capacity to hugely reduce the workload of radiologists. Therapixel, a French AI software company has developed an algorithm to interpret mammograms, called MammoScreen, which is now FDA approved and currently completing European regulatory requirements. Using deep learning technologies, the algorithm was trained to recognise cancer lesions on hundreds of thousands of previously confirmed cases. This will mean that the standard practice of two radiologists that inspects each mammogram can be replaced by one radiologist and the algorithm, which should reduce the time for a diagnosis.

In January 2020, Google Health published results showing their algorithm out-performed six radiologists, with fewer false positives and false negatives (McKinney, S.M., Sieniek, M., Godbole, V. et al. *Nature* 2020). Their DeepMind supercomputer was trained on de-identified mammograms from the UK and US and showed a 5.7−1.2% reduction in false positives when tested on new data and a 9.4−2.7% reduction in false negatives. "One thing that gives our AI system an edge is its large training set: over 70,000 mammograms, including more than 7,000 cancers. This is certainly more than a typical radiologist would encounter during their training," says Software Engineer and lead Google Health author Scott McKinney.

It is not clear why the AI model does better. "We're still investigating what perceptual features might drive this improvement," says McKinney. "We do know that the mistakes made by humans and the AI system are not perfectly aligned. For instance, when we showed cases to six independent radiologists, there were cancers that all radiologists missed, but the model caught. Conversely, there were also instances that all six radiologists saw, but the model missed." He sees this as a positive outcome that would allow for a superior combination of human and machine judgment. In August 2020, a study by researchers at the Karolinska Institute and University Hospital in Sweden that compared three different AI algorithms to identify breast cancer on previously taken mammograms demonstrated "that one of the three algorithms is significantly better than the others and that it equals the accuracy of the average radiologist."(Salim M, Wåhin E, Dembrower K, et al. *JAMA Oncol*. 2020)

Another AI model is being developed by Canon Medical Research Europe, in collaboration with Kevin Blyth of the University of Glasgow, to provide more accurate measures of tumour size in

mesothelioma, an aggressive cancer of the lung lining. “It grows around the lung like the rind of an orange, so it’s a very difficult tumour to measure,” says Blyth. The task is too time-consuming for radiologists who use much cruder size assessments and this can make it complicated to assess whether a patient has responded to a treatment. The hope is that AI will provide a solution. Blyth’s team has trained an algorithm to analyse CT scans and volume segment each individual image to find the total tumour volume. They are now ready to test the algorithm on new data, which will be collected as part of a pan-European mesothelioma research network funded by Cancer Research UK.

As noted in a recent article in Cancer World, the use of AI has been best displayed in diagnostic dermatology where AI outperformed expert dermatologists in diagnosing melanomas. However, diagnosis isn’t the only area where AI is providing support. In 2020, Steve Jiang from UT Southwestern Medical Center, used enhanced deep-learning models to create optimal radiology plans. As it is important to start radiation treatment as soon as possible in many patients, the ability to quickly translate complex clinical data into a radiology plan, could streamline the process. Jiang’s study showed an AI algorithm could instantly render 3D radiation dose distributions for each patient, trained on data from 70 prostate cancer patients, four AI models were able to predict the clinicians own calculations. UT Southwestern now plans to use these models with patients. (Nguyen D, McBeth R, Sadeghnejad Barkousaraie A, et al. *Med Phys*. 2020).

A team at IBM Research in Dublin are also looking at how AI can assist surgeons. “Our goal is to provide a surgical team with live tissue classification interactively, during surgery,” explains team member, Pol Mac Aonghusa, “think of it as a having a virtual pathology capabilities available during surgery to help the surgeons distinguish between tissue that is healthy, benign or cancerous.” The team has developed and trained an AI algorithm that can interpret subtle differences in the dynamic perfusion patterns of fluorescent dyes in real time (Zhuk S, Epperlein JP, Nair R et al. *Proc MICCAI2020*). “The indications from our early research is that the system can distinguish different tissue types with very high confidence – and in a matter of minutes,” so adds Aonghusa, ‘the potential is for more effective, less invasive surgeries with less post-operative complications.”

> ”The indications from our early research is that the system can distinguish different tissue types with very high confidence – and in a matter of minutes”

The AI genomics revolution

The biggest impact is likely to be seen in the interpretation of genomic data, according to Parker Moss, Chief Commercial & Partnerships Officer at Genomics England. So far AI is mainly being used in a ‘supervised learning’ capacity – that is, to do something a human does, but more efficiently, but says Moss, “I think that the more exciting area of machine learning and the much more disruptive area is ‘unsupervised learning’, where you have it look at complex data sets and don’t know what you’re looking for.”

Genomics England was set up by the UK National Institute for Health Research and NHS England and other medical charities to sequence 100,000 whole genomes that would give insight into rare diseases and common cancers. Using this data with novel AI algorithms they hope their collaborators will come up with patterns that might better predict the appropriate treatment or provide a more-accurate prognosis. “It is not just the 22,000 genes [in the genome], but it’s the complex gene networks,” says Moss, “that’s where machine learning comes in. [It’s] very good at

identifying signal in complex data and [the] whole genome is about the most complex data item you can get from an individual patient." Gerstung agrees we are reaching the era where we have sufficiently large genomic data sets to train complex algorithms to understand the unique mutational distributions along chromosomes that can be characterised as tumour subtypes.

One project Genomics England funds is a partnership with precision AI platform company Cambridge Cancer Genomics (CCG.ai) who are developing a sequencing panel that will cost-effectively profile the overall tumour mutation burden (TMB) – the total number of DNA mutations in cancer cells. Patients with high number of mutations are more likely to respond to certain immunotherapies. They hope this will allow them to assess DNA in the blood and use a 'liquid biopsy' approach, rather than having to repeat tumour biopsies and whole-genome sequencing.

Computational biologist, Moritz Gerstung, from EMBL's European Bioinformatics Institute (EMBL-EBI) in Cambridge, UK has been taking the image recognition used in analysing mammograms one step further and has found an algorithm that can infer genomic information from histology slides. "The algorithm that we used was originally developed by Google to recognize everyday objects on the internet that range from an Irish Setters to 'Spaghetti alla Carbonara'," he says. But now it is able to identify more than 160 recurrent DNA mutations, and thousands of RNA alternations in a tumour. Together with researchers from the Wellcome Sanger Institute and Addenbrookes Hospital in Cambridge, Gerstung has combined digital pathology with machine learning by training his algorithm on more then 17,000 digital histopathology slides covering 28 cancer types (Fu, Y., et al. *Nat Cancer* 2020).

"[The digitized slides were] a bit of a sleeping beauty, because, while people have extensively analysed all the different layers of molecular data, they hadn't really taken such a deep dive into the histopathological slides and how these could be related to genomic makeup," says Gerstung. The algorithm is able to analyse features such as nucleus size, size distribution and irregular positioning, at a statistical level – something not easily done by a human. "We found an association roughly for 20% of the genomic alterations," he says, but predicts with more data their algorithm would identify even more associations. Whilst this isn't a replacement for genetic testing Gerstung says there may be situations when this could act as a diagnostic tool and he hopes that it may be ready for the clinic within a couple of years.

Other groups are targeting specific genetic features, such as identifying microsatellite instability (MSI) in colorectal cancer, based on routine histology slides (Kather, J.N., Heij, L.R., Grabsch, H.I. et al. *Nat Cancer* 2020). "It would save the sequencing cost and even better, you could go through the back catalogue of patients who had already had a digital pathology image made of their tissue and may now be eligible for a drug which has comes to the market based on their MSI." says Moss.

IBM Watson Health, AI and genomics

IBM was an early player in AI, naming its supercomputer Watson after IBM's founder and promising to bring AI directly into the clinic to support oncologists with therapeutic decision making. It claimed that using natural language processing to extract information from peer-reviewed articles it could match genetic alterations in a patient's tumour with the most relevant therapies and clinical trials, matching or surpassing the clinical decisions made at top institutions.

But the path has not been smooth for IBM Watson and early indications suggested it would struggle to meet the expectations it had set. Reports in 2017, from a partnership with the Memorial Sloan−Kettering Cancer Center in the US, documented that Watson had difficulty distinguishing between cancer types and gave incorrect treatment options, such as suggesting patients with severe bleeding be treated in ways that would exacerbate the bleeding. Another US partnership with the

M.D. Anderson Cancer Center cost $62 million and was never actually tested on patients.

The past few years have shown signs of renewed confidence. In 2017, IBM and the University of North Carolina published the first paper on Watson's effectiveness, showing Watson spotted potentially important mutations not identified by a human review in 32% of cancer patients enrolled in the study (Patel, N.M., Michelini, V.V., Snell, et al. *The Oncol* 2018). In 2019, IBM signed a partnership with the University Hospital in Geneva (Hôpitaux universitaires de Genève, HUG) to use IBM's Watson, making it the first university hospital in Europe to use the tool, which IBM says completes an analysis of a whole genome and RNA-sequencing results in 10 minutes, compared to the 160 hours it would take to do this manually. Whether this new initiative will be successful remains to be seen. HUG did not respond to requests for information on the partnerships' progress and IBM Watson Health were unavailable for comment.

Challenges ahead

There is still scepticism around the clinical use of AI. One thorny issue is that of inbuilt biases. How the system learns will depend on the data used in its training. For example, when AI has been used to predict a person's age from an image, accuracy varies across ethnicities, unless the system is trained using racially diverse data sets. Google's model for analysing mammograms worked across data from the UK and US, but McKinney says they realise it may not be representative of all women around the world, "we're actively sourcing diverse new datasets to promote the inclusiveness of our technology," he adds.

The potential to exacerbate existing health inequalities in groups that are already under-served was pointed out in a recent report on AI for genomic medicine by the PHG Foundation, a Cambridge University think-tank. The other major issue with AI has become known as the 'black box' problem. The nature of deep learning algorithms means we often don't actually know how they produce their predictions – it's effectively a black box. "That's problematic," says Dr Philippa Brice, External Affairs Director at PHG, "although not insurmountable. AI still produces some major errors." In a talk at ETH Zurich in October 2017, Olivier Verscheure set out some of the problems still apparent in AI. Verscheure is head of the newly created Swiss Data Science Center, a joint venture between Swiss universities ETH Zurich and EPFL. He described how easily AI algorithms could be fooled. A recent image recognition test trained an AI system to recognise pictures of socks, but when only a few pixels of such an image were altered, the best algorithms identified the image as an Indian elephant. It certainly shows there is still a need for human supervision in areas such as cancer diagnosis, where it's important to understand the basis of any decision.

"I see this more as a timing challenge – rather than a fundamental incompatibility between AI and medicine," says Aonghusa, "It is fair to say that, like any new technology, a certain air of mystery surrounds the workings of some AI algorithms. But this is changing generally and it's fair to say that improving the explainability and trust in AI is one of the hottest topics in AI research." The new ambition is for 'white box AI' – interpretable models where we understand the variables that influence them. For example Aonghusa says their algorithm used to classify cancerous or healthy tissue is based on well understood physical values, with no hidden layers or complex parameters.

> "I see this more as a timing challenge – rather than a fundamental incompatibility between AI and medicine."

Combining human expertise with AI

For the time being AI is likely to be limited to specific data-heavy tasks, "which it will probably do better than a human," says Blyth, "but a human will still do more-complex tasks more naturally." So whilst his algorithm may more consistently analyse tumour volume in a mesothelioma patient, only a human can currently look at the same set of images and see that the disease has spread to other areas.

Ultimately Gerstung says "[we]should not think so much about AI, but rather about intelligence augmentation…at the end of the day, it will have to rely on an expert to make the definitive diagnosis, and I don't see that this will change anytime soon." But it certainly has the capacity to free physicians from administrative, clerical tasks so they can focus on the uniquely human work of connecting with patients.

What AI may stimulate is a new type of clinician. "I think it is changing the face of medicine," says Brice, "I think we do need new supporting health professions, scientists who can explain and interpret data to aid the clinicians." But Aonghusa says he can't see AI fundamentally changing things, "the role of the physician has continued to evolve for hundreds of years. It has continuously adopted and applied new technologies from germ theory to antibiotics – AI should be no different." He says from his experience of working with trainee surgeons the adaptation to AI and the opportunities it promises is already occurring.

Nevertheless AI is likely to increase the push towards more interdisciplinary medical teams and for Genomics England a more joined-up approach to clinical practice and clinical research. "For machine learning to have an impact, the two worlds of research and clinical care have to come together. That's very much what Genomics England are trying to do," says Moss. He describes a "virtuous infinity loop," where the patient consent to use more data as a resource for machine learning will drive diagnostic and therapeutic advances that will then encourage more clinicians to use it.

> "That's where the problem comes…it's a good servant, it's a great servant, but a bad master."

Obviously the use of massive amounts of patient data brings up the issue of ethics and consent. Genomics England stress their debt to the 100,000 participants who have agreed to contribute their detailed medical records, albeit anonymised. "[We] do everything we can to reassure participants, that their data is being used for the right purposes in a safe and secure environment," reassures Moss. This includes communicating to patients how AI works, which he says "can be conceptually difficult to understand."

AI and machine learning offer unique possibilities in oncology but says Brice "the danger we have to resist all the time, with all these technological advances is to think it's the solution to everything and we don't need to think about how to use it. That's where the problem comes…it's a good servant, it's a great servant, but a bad master."