



DANIEL STOLLE

Deep learning delivers early detection

Artificial-intelligence systems look set to make early detection of lung cancer more accurate and widely available. **By Elizabeth Svoboda**

After years of helping to train an artificial-intelligence (AI) system to find the early stages of lung cancer, Mozziyar Etemadi was thrilled when the computer found tumours in scans of patients more accurately than trained radiologists did¹. He was even more excited when his team gave the system old computerized tomography (CT) scans of the chests of people who later developed lung cancer. No doctor had seen anything amiss in these early scans, but the machine did.

“A human would say this was normal,” says Etemadi, a biomedical engineer at Northwestern University’s Feinberg School of Medicine

in Chicago, Illinois. “But the AI was discovering these subtle patterns, and it was very confident. It was finding the cancer.” After the machine completed a run, Etemadi thought: “We just discovered this guy’s lung cancer a year or two before we would have otherwise.” His mind raced at the prospect of boosting the survival chances of thousands of people.

Lung cancer is the deadliest cancer in the world – about 75% of those who have it die within five years of diagnosis. But when cancers are found early, the prognosis is much better. If tumours are small and confined to the lung, almost two-thirds of people survive for at least five years.

The need for early detection has fuelled the development of AI systems that can detect ever-smaller lung tumours. The system Etemadi is working on – a joint initiative between Google, Northwestern University and other institutions – is one of several now moving towards clinical adoption. In July 2020, the University of Oxford, UK, announced an £11-million (US\$14.3-million) research programme to use AI to help diagnose lung cancer.

Such developments promise to make lung-cancer screening more precise and accessible to all. But turning the new systems into clinical mainstays will require careful cultivation of the relationship between radiologists and the machines on which they depend.

Spot the tumour

About 70% of lung cancers are detected in the later stages of the disease when it is harder to treat, which partly explains why the 5-year survival rate is so low. Initial symptoms of lung cancer tend to be common maladies, such as a persistent cough or fatigue, which are easy to dismiss as inconsequential. “People ignore a cough,” says oncologist Mariam Jamal-Hanjani at University College London’s Cancer Institute. “People often come to my clinic with metastatic disease,” she says, but by

that stage effective treatment might already be out of reach.

Studies at the University of California, Los Angeles, and elsewhere show that regular screening of at-risk populations can detect many cases of lung cancer much earlier, reducing mortality by 20–30%. The US Preventive Services Task Force, a volunteer group that makes recommendations for clinical preventative services, now recommends annual CT screening in groups at high risk of lung cancer, such as past or current smokers.

But the number of radiologists who assess lung scans has not increased enough to keep up with rising demand. “There are so many CT scans, so many people,” says Ulas Bagci, a computer-vision specialist at the University of Central Florida in Orlando. This intense workload can cause overstretched radiologists to make mistakes.

The limits of human vision also make it easy for radiologists to overlook tiny malignant lesions. Up to 35% of lung nodules are missed at the initial screening, for example. Using AI systems can help on both counts by shifting some of the burden from busy specialists and detecting lung spots invisible to the naked eye.

Radiologists already use computer-aided diagnostic tools to help them to spot malignant tumours. Typically, a human programmer tells the system what features to look for but the computers flag a lot of presumed malignancies that are actually benign. “The radiologists didn’t like it because they’d need to click each of them” to check, which wasted a lot of time, Bagci says.

More recent AI systems are based on a principle called deep learning. Rather than looking for tumour features defined in advance by a programmer, deep-learning systems figure out for themselves what a tumour is from real-world examples. Researchers give the systems a large data set comprising thousands of people’s lung CT scans, some with cancer and some without. From this, the machines learn for themselves what a lung cancer nodule looks like.

The more training scans the systems view, the more reliably they can distinguish lung tumours from benign splotches. And they do so more accurately than older, non-AI systems. Some of the deep-learning systems also give clinicians an estimate of how confident they are in their judgement, which can further inform clinical decision-making.

Going deeper

Etemadi’s system relies on this deep-learning approach to identify lung tumours on CT scans. In 2019, he and his team reported that their system correctly identified the early

stages of lung cancer 94% of the time, outperforming a panel of 6 veteran radiologists¹.

The researchers trained the system using a database of more than 40,000 CT scans – not just current ones, but also scans from before people received a lung cancer diagnosis. During this training period, the scientists told the computer which early-stage scans turned out to contain cancerous spots and which did not. Over time, the computer learnt which image properties separated malignant spots from benign ones, and it became better and better at flagging early signs of cancer.

The system’s ability to analyse an entire 3D CT scan, rather than just a sequence of 2D slices, also improves its accuracy. What’s more, 3D scans provide more diagnostic information about features such as blood vessels that are not part of the main tumour, Etemadi says. “The 3D volume starts highlighting areas far away from the tumour. It has shown us some things we wouldn’t expect. We’re opening up a whole new area of scientific enquiry.”

Bagci and his team have developed another deep-learning AI model that is similarly skilled at detecting nodules that indicate early lung cancer. The computer correctly identified tiny specks of cancer on CT scans about 95% of the time – much higher than the 65% accuracy rate that radiologists typically achieve.

“It has shown us some things we wouldn’t expect. We’re opening up a whole new area of scientific enquiry.”

Both Bagci’s and Etemadi’s systems view the scans multiple times. First, they scan for irregular areas such as oddly shaped blotches that might be cancerous. Then, they assess each of these target areas in more detail to make a final judgement about whether they are malignant.

Bagci’s team trained their system on CT scans containing tumours around 1–3 millimetres in size, which many radiologists find hard to spot. “It’s very difficult to visually search all the pixels on the screen. There’s a huge rate of missing those,” Bagci says. Because his AI system is trained on thousands of lung scans, it is highly optimized to detect tiny problem areas that specialists might overlook, he says. “You are able to use more data and more powerful algorithms. It started finding the small nodules better.”

Another deep-learning system, developed by Jamal-Hanjani’s group at University College London and London’s Institute for Cancer Research, tackles the related problem of detecting early signs of lung cancer recurrence after

initial treatment. The team reported² this year that after they trained the computer on hundreds of images of early stage lung tumours, the system figured out that tumours with regions low in immune cells are more likely to trigger a relapse after surgical resection or chemotherapy. The scientists think this is because these tumours have some form of cloaking mechanism to evade the immune system, allowing cells to divide unchecked. These warnings of a potential relapse could help radiologists to identify people who need careful monitoring, Jamal-Hanjani says.

Wide screen

One of the main advantages of deep-learning systems is that they can speed the advent of population screening, which is aimed at catching lung cancer much earlier. There is strong evidence that such programmes would be effective.

Researchers at Erasmus University Medical Center in Rotterdam, the Netherlands, for example, recently studied³ the impact of a screening-programme trial in Belgium and the Netherlands. The team followed more than 15,000 current or former smokers over 50 years of age for at least 10 years. By the end of the trial, people who had undergone regular screening were about 25% less likely to die of lung cancer than were the controls – and only 1.2% of participants had a false-positive scan, which indicated cancer where it was not actually present.

A US actuarial analysis⁴ found that a national lung cancer screening programme for high-risk individuals would cost about \$19,000 per life-year saved. This compares favourably with existing screening programmes for breast, cervical and colorectal cancers.

With AI doing some of the heavy lifting, lung cancer screening programmes could prevent similar numbers of deaths at an even lower cost thanks to increased automation and without imposing as much of a burden on radiologists. Daniel Tse, a product manager at Google Health, says that AI-assisted screening could help to identify not just people with early-stage lung cancer, but also those who are at high risk of developing lung cancer within the next few years. AI is “not a panacea” to simplify broad-based screening, Tse says, “but we think it can be a very powerful tool”.

Jamal-Hanjani says radiologists will soon be able to combine screening results with genetic data to create even more customized treatment plans. As deep-learning systems churn through different kinds of large data sets, such as CT scans, genetic sequences and treatment histories, they often discover unexpected relationships. For instance, a pattern that shows up



Mozziyar Etemadi develops artificial-intelligence systems that find early stage lung cancer.

on someone's CT scan might predict that the tumour will have a particular genetic make-up. A clinician could follow this up by sequencing their tumour cells to see whether this prediction is correct. This could help the care team to choose the most appropriate type of treatment for that specific cancer variety.

An evolving partnership

Before scenarios like these can become a reality, doctors and AI researchers need to address urgent questions about how best to interpret the results that computers find – and how to divide the diagnostic workload between machines and trained physicians.

That deep-learning systems can outperform humans on some diagnostic tasks does not mean that they will take over radiologists' jobs. Deep-learning systems can supply diagnostic guidance, engineers say, but they cannot yet replace human specialists. They are likely to enhance physicians' diagnostic skills, rather than make them obsolete, says Bagci. "Computers are good at very local tasks. Humans are much better at global tasks," such as rendering a definitive diagnosis from multiple information sources, including blood tests and physical examinations as well as scans, he explains.

Tse agrees with this assessment, adding that humans are better at learning quickly about the minutiae of unusual lung cancer cases. AI systems, on the other hand, excel at flagging common types of early cancerous lesion, having been trained on data sets that include thousands of such cases. "The majority of cases that a doctor sees are bread and butter," Tse says. "That's where we want to assist. We want to help people be more efficient with their time."

There are issues of trust to be overcome if

radiologists and AI are to work in harmony. By tapping into previously unimaginable stores of computing power, AI systems can now assess millions of different variables in a single scan before rendering a judgement of, say, "almost certainly benign" or "75% chance of malignancy". But the more complex the image analysis, the harder it is for the system to describe what it is doing in ways that humans can understand. Millions of equations, Bagci says, do not translate easily into explanations of why a particular diagnostic recommendation has been made. "These algorithms are really a black box. Why cancer? It doesn't tell you."

"Two brains are better than one. It's going to make radiologists' and other physicians' jobs easier."

Researchers are starting to devise diagnostic systems that provide clearer explanations of their advice. Bagci, along with a team at the US National Institutes of Health, has recruited radiologists to help develop a new type of deep-learning system. When they train the software, researchers use an eye-tracking device to capture how the specialists analyse each scan. In initial tests, this radiologist-trained system is proving more than 90% accurate⁵ in detecting cancerous spots. "The AI learns where the radiologists look," Bagci says.

Training deep-learning systems with such input, as well as with radiologists' own analyses of why certain spots look like cancer, can help the systems to draw on more-transparent

and understandable criteria for their results. To justify a recommendation, for instance, a system might point out that a spot has a wavy border or that a lesion's characteristics have changed since the previous scan.

Gradual progression

The difficulty of integrating AI into radiologists' workflow leads Etemadi to envisage lung cancer diagnosis becoming more automated in a series of small, gradual steps. A computer system could initially do a baseline reading of each lung scan and present them in the order the doctor prefers – from easiest to hardest to interpret, for instance, or from the highest to the lowest likelihood of lung cancer.

The advantage of adopting AI incrementally is that radiologists will not have to suddenly change the way they work. Deep-learning tools can be incorporated into the existing computer-aided diagnosis systems, says engineer Andrew Berlin at the Draper Laboratory in Cambridge, Massachusetts.

As clinicians and engineers feed ever more lung CT scans into deep-learning systems, privacy must remain paramount, says Andrew Crawford, policy counsel for the Center for Democracy and Technology in Washington, DC. "I am all in favour of using technology to drive better patient outcomes, as long as you're doing it in a way that's going to engender trust," he says.

Etemadi points out that Google and many AI-research institutions already have some privacy measures in place, such as removing names from lung-scan data and obscuring the date each scan was taken.

If developers can navigate the very human concerns that go along with outsourcing some precision tasks to AI systems, lung cancer specialists who try the new tools will not want to go back, Bagci says. "Two brains are better than one. It's going to make radiologists' and other physicians' jobs easier."

And that will result in better long-term outcomes. Using AI to find tumours early can effectively double the amount of time oncologists have to treat a patient, giving them much more opportunity to keep the cancer from spreading. As Etemadi says: "Most patients who get lung cancer end up dying. You have the potential to really change that fraction."

Elizabeth Svoboda is a science writer in San Jose, California.

1. Ardila, D. et al. *Nature Med.* **25**, 954–961 (2019).
2. AbduUabbar, K. et al. *Nature Med.* **26**, 1054–1062 (2020).
3. de Koning, H. J. et al. *N. Engl. J. Med.* **382**, 503–513 (2020).
4. Pyenson, B. S., Sander, M. S., Jiang, Y., Kahn, H. & Mulshine, J. L. *Health Affairs* **31**, 770–779 (2012).
5. Khosravan, N. et al. *Med. Image Anal.* **51**, 101–115 (2019).