Unsplash © National Cancer Institute

The hard truth about AI’s ability to help us beat cancer

Back in 2016, a series of bold headlines claimed that artificial intelligence had reached the level of "boosting cancer screens to nearly 100% accuracy." By 2018, AI was all set to make "doctors obsolete."

In the years since, AI has been repeatedly labelled as the silver bullet in our fight against breast cancer and, with every announcement, it feels like there may be light at the end of the tunnel. Yet, at the same time, the numbers of cases and deaths have continued to rise.

As of 2020 – the last year for which we have full and verified figures – breast cancer cases made up 22% of all cancer cases globally – roughly 4M people. In the same year, breast cancer *overtook lung cancer* to become the most commonly diagnosed form of the disease worldwide. It now affects one in eight women globally and estimates suggest that, even with all the recent technological advancements being made in this field, the number of people being diagnosed with cancer will rise by as much as 50% in the coming years.

This is despite the fact breast cancer is one of the most well-researched, well-funded and well-understood cancers of our time. A cancer which, when caught early, comes with a 90% five-year survival rate. That is, five years after diagnosis, 90% of people will still be alive.

There is clearly a disconnect somewhere in the chain. The potential and promise being made in labs are not filtering through to clinics. The numbers suggest millions are dying from a disease that, at least statistically, they should survive. So what's holding it back?

AI's potential and pitfalls

A person's chances of surviving breast cancer typically correlates with how soon their cancer is caught. At early stages, surgery is commonly used to remove the tumour(s) without the need for systemic treatments like chemotherapy, or a full mastectomy. That's not to say later stage cancers can't be treated effectively, or that early-stage cancers don't prove terminal, but the chances are smaller.

At the heart of catching breast cancers early is screening. This is where the majority of advancements, and headlines, about AI's impact have been concentrated. *Evidence* shows that when women attend "organised, population-based mammography screening programmes," breast cancer mortality is reduced by around 20%, compared to groups of women who don't attend. This is seen across all age groups, not just those in the highest risk categories. Further to this, a Swedish *study* found that women who participated in mammography screening had a 41% reduction in their risk of dying of breast cancer within 10 years.

The processes involved in breast cancer screening lend themselves almost perfectly to the capabilities of AI and machine learning. The quality of a screening programme is rated by how well it detects *cancers*, referred to as its sensitivity, as well as its recall rates and positive predictive values, or its specificity. These two measures can be quantitatively tracked and verified, which is the bread and butter of data science, and hence AI.

> *AI can be trained on millions of images and can identify patterns that may, to the naked eye, appear hidden.*

The screening workflow similarly suits AI's capabilities. Under the NHS Breast Screening Programme (BSP), every mammogram is analysed by two separate readers who, in the majority of cases, are trained radiologists. If these readers don't agree on the required outcome for the patient, the mammogram is sent to arbitration.

Such due diligence is necessary because it reduces the rates of missed cancers. It also improves the number of false positives, both of which can cause psychological and physical damage to patients. Yet such due diligence is time-consuming, expensive and requires a large amount of resources.

In this way, AI can be inserted at any one of these steps in a clinical workflow to save time, money and resources. It can lighten the load of a radiologist by complementing or replacing readings. It can be used to spot anomalies. AI could filter out so-called “normal” readings, referring only those with anomalies to the human readers, or, when paired with automation, can take on the heavy lifting of admin and reports, leaving radiologists with more time and resources for analysing suspicious readings and patient care.

Unsplash © Artem Beliaikin

When AI just isn’t enough

There is, however, a catch. In fact, there are two. Firstly, there are significant liability, regulatory and cultural hurdles to overcome when introducing AI, or any new technologies, into clinical routines. Secondly, research has increasingly shown that AI, when used alone, just isn’t enough.

“While it’s a new and exciting area, almost anything that provides diagnostics using AI, particularly using image processing, is problematic from regulatory and adoption points of view,” says Peter Finnie, partner and patent attorney at Potter Clarkson.

“Firstly, there’s the issue of getting patentability. Generally speaking, the process requires the right timing and understanding to ensure intellectual property is protected correctly. With digital innovation not fully recognised by the UK patent office, the intricacy of AI complicates this further, as does its fusion with the guarded medical diagnosis process.

“Next, there’s regulatory consideration. The introduction of machines for health-related decisions brings up the question of ‘Who’s liable?’”

> *There are also cultural barriers to navigate.*

“To get technology such as predictive AIs into clinical settings where they can save lives means tackling negative connotations and fears," explains Owen Johnson, senior fellow at Leeds University. "AI challenges people and their attitudes: the professionals that the machine can outperform, and the patients that are reduced to statistical probabilities to be fed into complex algorithms. Innovation in healthcare can take decades."

Finnie adds: "Doctors don't know how an AI-based healthtech invention has been trained or the reliability at the edges of its abilities to make predictions. And while the software may have regulatory approval, there's going to be a natural reticence over whether it should be added into the clinical decision pathway they've crafted over a 30-year career."

Elsewhere, as part of a scoping *review* led by Nehmat Houssami from The University of Sydney's Faculty of Medicine and Health in 2019, researchers found that while AI models generally report "good accuracy" for breast cancer detection, "methodological concerns, and evidence gaps" exist that prevent these successes being adopted in clinical settings. These concerns centre on the fact that studies are largely retrospective in nature, and that they rely on small and highly selected image datasets. Some of the methodologies didn't adequately account for the potential of bias in AI model training, while others lacked comparative data.

In September, researchers led by Karoline Freeman from the University of Warwick Medical School went a step further and concluded that AI, on its own, is simply not ready yet. Her team analysed 12 studies totalling 131,800 screened women and, like Houssami, described the methodological quality across the range as lacking. In fact, she described it as "poor."

"Current evidence on the use of AI systems in breast cancer screening is a long way from having the quality and quantity required for its implementation into clinical practice," Freeman explains, citing the fact that in three of the studies, the AI screened out as much as 10% of women who had cancer. Cancer that was successfully identified by human radiologists.

Unsplash © National Cancer Institute

If you want something different, you have to do something different

In order for AI to be effectively integrated into breast cancer screening programmes, Freeman claims there will need to be "well-designed comparative test accuracy studies, randomised controlled trials, and cohort studies in large screening populations which evaluate commercially available AI systems in combination with radiologists."

In fact, the latter of Freeman's recommendation is key to directly addressing both of the catches in getting AI into clinical use. Instead of focusing on whether AI can (and should) replace radiologists, its potential should be shown in how well it complements its human colleagues. To allay fears, appease regulators, and take AI's potential beyond the technology alone.

This is likely why a number of startups are now taking this approach, including German deeptech firm *Vara*. Vara has built a solution to the AI in breast cancer screening dilemma from the ground up. It set its overall mission – to catch every deadly cancer early – and built an AI-powered, full-stack platform to achieve it.

"History books are full of AI models that worked retrospectively but failed to show impact prospectively in new settings," according to Jonas Muff, Vara's CEO. "It's not enough to base an AI tool or platform's performance on an average score. Women and their cancers rarely fit neatly into the 'average' conditions and thus AI on its own fails a wide spectrum of subgroups. Instead of an AI plug-in, we've built an end-to-end platform that reimagines the entire clinical workflow.

During development, Vara analysed every step in the breast cancer screening journey to identify the most pressing pain points, almost always in collaboration with radiologists themselves. Its platform attempts to address each of these individual pain points in order to improve the journey as a whole. It uses AI and automation together, but these advanced technologies form just one part of the platform's appeal. Instead of replacing radiologists, it is set up to run alongside them. Instead of retrospectively analysing data, it takes a prospective route.

Vara is one of a small group of tools and platforms that are already in use in clinical settings. It has analysed more than 7M mammograms and is used by 30% of screening units in Germany. Earlier this year, Vara additionally launched screening units in Mexico, and Greece in partnership with healthcare providers on the ground in those regions.

"The Vara platform captures real-world data in clinical use. This includes monitoring AI continuously while providing screening radiologists with key performance metrics. This is an invaluable asset," continues Muff. "For every woman diagnosed via Vara, we can follow the patient pathway beyond screening to truly assess our AI's impact on population health metrics. This creates a good deal of transparency as to how the AI works and how it's impacting screening performance, which in turn helps create a data feedback loop from prospective clinical use, allowing us to constantly improve our AI models"

Taking a more global approach

It’s outside of Germany, through Vara’s partnerships in Mexico and Greece, that talk more to the true, global potential of AI in breast cancer screening generally.

The 90%, five-year survival rate mentioned above only applies to high-income countries. In India, only 66% of patients fight off the disease. In South Africa, this drops as low as 40%. It also just so happens that these countries, alongside many more, don’t currently have national, organised and effective screening programmes. There is a widescale and devastating diagnostic gap that results in millions of women being left behind without access to the potentially life-saving tools their peers in other countries often get for free.

> *Just as there is a correlation between breast cancer screening and breast cancer survival, there is a correlation between a lack of screening and increased mortality.*

While competing with the same liability, regulatory and perception hurdles as high-income regions, these countries have their own, nuanced barriers to break through. They don’t typically have the money to invest in screening infrastructure. They don’t have the technology to power it and they don’t have the resources, including the required number of trained radiologists to drive these programmes.

This is where AI can come into its own. It can reduce the costs involved in running screening programmes, support the lower numbers of trained radiologists, and help health systems better manage their scarce resources.

“Most modern AI solutions promise to reduce costs for those healthcare providers and systems that already have the infrastructure in place – those that are already running screening at high volumes,” Muff adds. “Yet technology’s power lies in democratising access to such benefits. Increasing access to effective breast cancer screening to women should be the starting point, not an afterthought.”

Vara’s platform, as part of its “catching deadly cancers early” mission, was built for such scenarios. It’s why, Muff tells us, that it’s a cloud-based tool. This design means the platform can effectively be accessed anywhere with an internet connection. Even in regions that don’t offer routine screening, and that don’t have the level of wider infrastructure seen in regions that do.

The platform helps the scarce number of radiologists work more efficiently and, via pop-up screening units, Vara’s partnerships with local healthcare systems allows mammography scanners to be installed wherever women need them.

“If something significant is not done now, annual breast cancer deaths are expected to exceed a million in 2040,” concludes Muff.

Just as AI can’t beat breast cancer on its own, neither can the startups working in this space. The complex network of regulation, liability and perception is what’s holding AI back from its true potential. It’s going to take a combined effort from governments, clinicians, and patients to finally unlock the technology’s power if we’re going to be able to use it to prevent millions of unnecessary deaths.

Article by VICKY WOOLLASTON