The Untapped Social Impact of Artificial Intelligence for Breast Cancer Screening in Developing Countries: A Critical Commentary of DeepMind

Joe Logan,[1,2] Paul J. Kennedy,[2] Daniel Catchpoole[2,3]

[1]Alixir Technologies Pty Ltd, Sydney, New South Wales, Australia
[2]Faculty of Engineering and Information Technology Sydney, Sydney, New South Wales, Australia
[3]The Tumour Bank, The Children's Cancer Research Unit, Kids Research, The Children's Hospital at Westmead, Sydney, New South Wales, Australia

Address correspondence to Joe Logan (joe@alixir.ai).

© Innovative Healthcare Institute

INTRODUCTION

In January 2020, Google's DeepMind team published an article demonstrating that a deep neural network–based artificial intelligence (AI) system could outperform a human radiologist at the task of interpreting mammograms.[1] The most ground-breaking aspect of this study is not the machine learning architecture itself, rather it is the fact that the authors trained their system by using a wholly histologically labelled dataset, rather than data that used the radiologist's opinion as the ground truth. With DeepMind focusing exclusively on British and American patients, this commentary discusses how they may have missed the social impact use-case for the technology to address the needs of the 5 billion women who do not undergo breast cancer screening in the developing world.

The incidence of breast cancer is rapidly growing in the developing world, regions that do not have the financial or human resources to implement a traditional radiologist-led screening program. In these circumstances, the scalability and low cost of AI systems, such as that put forward by DeepMind, could be a viable solution. According to the current position statement of the World Health Organization (WHO)[2]:

- Breast cancer is the most frequent cancer in women worldwide and is increasing, particularly in developing countries where most cases are diagnosed in late stages.
- Breast cancer survival rates vary greatly, ranging from 80% in North America, Sweden, and Japan to around 60% in middle-income countries and below 40% in low-income countries.
- The low survival rates in less developed countries can be explained by the lack of screening.

Scalable cost-effective technology, such as that put forward by DeepMind, has the potential to redefine healthcare in the developing world, and prevent thousands of lost lives. This commentary provides a critique of DeepMind's work, in the context of the true scale of the breast cancer problem in the developing world, making a strong case for further refinement and implementation of the technology into locations that would benefit from it the most.

THE IMPORTANCE OF BREAST CANCER SCREENING

Twenty-two countries run large-scale government-funded breast cancer screening programs, including the United Kingdom (UK) and Australia.[3] The importance of breast cancer screening and early detection is widely acknowledged, and is associated with a doubling in 5-year survival rates when compared to countries that do not screen women for breast cancer.[4] The reasons behind this are self-evident. Screening is levelled at detecting a problem before it becomes too big to manage, and breast cancer is a prime example of a cancer that can be completely cured if it is detected early.

In the countries that do provide breast cancer screening, the format and process behind it is largely the same. Mammography (two-dimensional breast x-

rays) is the chosen imaging modality, driven largely by the low cost and speed by which the test can be administered. The resultant scans are then interpreted by up to two radiologists. It is generally understood that the provision of the radiologists represents the largest variable cost to screening programs worldwide.[5]

The WHO's global statistics show an alarming increase in the incidence of breast cancer in the underdeveloped regions, which they attributed to increased westernization (including alcohol use, smoking, and obesity) affecting the developing world.[2] Unfortunately, as is the case for the most developing countries, there are simply not enough clinical or financial resources to implement viable screening programs that prevent high mortality rates.

THE DEEPMIND STUDY

The groundbreaking aspect of DeepMind's work in mammography was the team's choice to use large-scale wholly histologically confirmed training data.[1] Traditionally, researchers have been forced to rely on clinically subpar datasets that are either too small for real-world use, or are labelled on the radiologist's opinion rather than histologic confirmation.[6] DeepMind had access to 28,953 histologically confirmed mammograms from the OPTIMAM dataset,[7] in addition to 3097 supplemental mammograms from the United States (US). Through applying an end-to-end deep learning model—the architecture of which was not disclosed—to this dataset, DeepMind was able to demonstrate the following results when compared to a cohort of radiologists in both the UK and the US:

- An absolute reduction in false positives between 1.2% and 5.7%.
- An absolute reduction in false negatives between 2.7% and 9.4%.
- An overall AUC-ROC (area under the curve-receiver operator characteristics) improvement of 12.1%.

In an interesting footnote in the DeepMind article the authors compared their results on the US demographic; they used a model trained on UK data alone and another model that also incorporated the 3097 supplemental US mammograms. They demonstrated that supplementing with US data led to an overall improvement in their AUC-ROC score of 7.09%. This highlights the importance of incorporating data from the target countries to extract the best possible results.

An obvious pitfall of DeepMind's research is that their model was only validated and measured in largely Caucasian countries. Both countries already run world-class breast cancer screening services, albeit with some major differences in how they are funded. Suffice it to say that breast cancer care in both the United Kingdom and the United States is first-world, with comparatively high detection rates and 5-year survival rates of approximately 80%.

BREAST CANCER STATISTICS IN THE DEVELOPING WORLD

According to the WHO, the mortality rate from breast cancer is over twice as high in the developing world relative to the developed world, where established screening programs are in place.[2] In their latest breast cancer statistics survey, the WHO made the following statement:

> "Early detection in order to improve breast cancer outcome and survival remains the cornerstone of breast cancer control."[8]

A 2010 meta-analysis on breast cancer trends in low- to middle-income countries concluded that 45% of the 1.35 million new diagnoses would occur in the developing world, which represents 55% of the world's breast cancer deaths.[9] The authors predict that in 2020, the breast cancer incidence globally will rise by 26%, which they attributed to the increasing rates in underdeveloped regions. The increasing rates in these areas are generally considered to be due to an increase in "Western" behaviors, which include smoking, alcohol, obesity, and delayed parity.[10] In India, it has been shown that the breast cancer incidence is rising at a rate of between 0.5% and 2% every single year.[11]

A study performed in Pakistan gives further weight to the impact of the growth of breast cancer in the developing world, by demonstrating that most patients clinically present at advanced stages (III and IV) rather than earlier when the disease is treatable.[12] This study showed that 65.7% of patients present in the advanced stages, compared to the 43.6% in middle- and high-income countries. The authors commented that the lack of an established screening program in Pakistan was the causative factor.

The growth of breast cancer in the developing world often presents a challenge for governments. For example, in India, with a population of 1.35 billion people, with a significant number living in poverty, there simply are not enough human or financial resources to implement a typical radiologist-led screening program.[4] However, if the requirement for radiologists within the screening process can be replaced by scalable, cheap, and available technology, then this could present an opportunity to tackle breast cancer for these impoverished areas of the world.

THE SOCIAL IMPACT OF USING AI TO SCREEN FOR BREAST CANCER IN THE DEVELOPING WORLD

DeepMind's article was positioned largely as the use of an AI system to augment traditional radiologist-led breast cancer screening. By demonstrating a median reduction in missed cancers of 6.05%, DeepMind's technology has the potential to detect cancer in 2541 women that would otherwise be missed in the US.

Although this is certainly not insignificant, the true impact of such technology is far more profound in the developing world. In India, it has been demonstrated that 131,000 women will be diagnosed with breast cancer in 2020 and approximately 52,400 will die as a result of it.[11] Based on the study by Coleman et al,[4] the 5-year survival rate from breast cancer is nearly halved in countries with established screening protocols. Therefore, it can be postulated that technology such as DeepMind's has the potential to save 26,200 women every single year.

At this juncture, it is appropriate to come back to the questions of why and how an AI system could achieve this degree of social impact, particularly over implementation ofradiologist-led breast cancer screening programs. Put simply, scalability and cost. If we reflect on the disease burden in India, it is clear that incorporating a radiologist-led program for a population of 1.35 billion people becomes economically challenging. Among this population, 70 million people would be eligible for screening, and taking a traditional radiologist double-read model and the lower end of the pricing tier at $10 (US Dollars) per read, costs could quickly exceed $1.4 billion per year for radiologists alone. To further compound this situation, there would be huge drain on the Indian radiologist workforce to maintain this model because 200,000 mammograms would need to be analyzed every single day. With a large number of Indian radiologists already contributing to Western screening programs through teleradiology, there simply are not enough human resources to make this approach viable.

However, if we replace the need for radiologists entirely by adopting AI reading, such as that put forward by DeepMind, then the viability of a screening program immediately appears achievable. Based on the cost and availability of GPU (graphics processing unit) cloud computing resources, or on-chip solutions such as Nvidia's Jetson,[13] we can determine that reading a mammogram could theoretically be performed by AI at approximately $0.001 rather than the $10 commanded by radiologists. This would reduce the interpretation cost burden to $70,000 per year, a 10,000-fold cost reduction. The true social impact of such technology becomes more attributed to its cost at scale rather than inherent accuracy.

DISCUSSION

DeepMind's research[1] is vitally important as a seminal piece of research that could pave the way to better adoption of AI in healthcare. This was driven exclusively by the team's choice of datasets that exclusively use a histologic ground truth, rather than one based on human opinion. This is the primary explanation for DeepMind's ability to train an AI system that can finally exceed the performance of a human radiologist on digital mammographic scans.

An obvious criticism of DeepMind's work is that the authors failed to acknowledge the potential of their technology in the battle against breast cancer in less developed geographic areas. A number of studies have highlighted the worrying growth of breast cancer in the developing world, including those of Coleman et al,[4] *The Lancet Oncology*,[11] WHO,[2] and Coughlin and Ekwueme.[14] For these regions, where population and budgets make traditional radiologist-led screening impossible, technology such as DeepMind's has tremendous potential to dramatically reduce the disease burden of breast cancer through its low cost and availability.

Although the potential of such technology to reduce the global disease burden of breast cancer is significant, there are a number of pitfalls in DeepMind's research that will hinder it from being universally adopted as a true replacement for a radiologist in the developing world. First is that the DeepMind team failed to disclose the technical architecture of their AI system. This inhibits future researchers from replicating their results and using it as a platform for further research and development. The second criticism is that DeepMind chose to exclusively use datasets from predominantly Caucasian patients. Although DeepMind made a strong case for the use of supplementary local data, clearly any attempt to create models that generalize globally require supplementary data from a number of diverse ethnic regions. Moreover, any models that are designed to generalize globally and be used in the developing world would require testing and validation in a variety of ethnogeographic regions. Another criticism and area for future research would be the use of film-based mammograms in the training data, in addition to digital mammograms. As many of the poorer regions of the world only have access to film-based mammographic scanners, it is likely that the inclusion of film mammograms in the training dataset would be mandatory.

In summary, DeepMind has provided a rationale for the importance of high-quality training data in medical AI, which, in turn, has further justified the ability of such technology to not just match, but exceed, the performance of a radiologist. Although this is impressive, no attention has been given to the real social impact that the scalability and low cost of this technology could have in developing regions of the world, where there simply are not enough resources to implement a traditional screening program. However, within this research lies a huge untapped potential, and if an AI system can be truly generalizable and provide better patient outcomes in areas where there is immense disease burden, we will certainly move the needle in terms of social change and medicine.

References

1. McKinney SM, Sieniek M, Godbole V, et al. International evaluation of an AI system for breast cancer screening *Nature*. 2020;577:89–94.

2. World Health Organization. WHO position paper on mammography screening. 2014. Accessed Jul 12, 2020. www.who.int/cancer/publications/mammography_screening/en/

3. Shapiro S, Coleman EA, Broeders M, et al. Breast cancer screening programmes in 22 countries: current policies, administration and guidelines *Int J Epidemiol.* 1998;27:735–742.

4. Coleman MP, Quaresma M, Berrino F, et al. Cancer survival in five continents: a worldwide population-based study (CONCORD) *Lancet Oncol.* 2008;9:730–756.

5. Brown J, Bryan S, Warren R. Mammography screening: an incremental cost effectiveness analysis of double versus single reading of mammograms *BMJ.* 1996;312:809–812.

6. Moreira IC, Amaral I, Domingues I, et al. Inbreast: toward a full-field digital mammographic database *Acad Radiol.* 2012;19:236–248.

7. Halling-Brown MD, Warren LM, Ward D, et al. OPTIMAM Mammography Image Database: a large scale resource of mammography images and clinical data. Posted online Apr 9, 2020. arXiv:2004.04742

8. World Health Organization. World Health Organization ongoing statement on breast cancer. 2020. Accessed July 10, 2020. www.who.int/cancer/prevention/diagnosis-screening/breast-cancer/en/

9. Tfayli A, Temraz S, Abou Mrad R, Shamseddine A. Breast cancer in low-and middle-income countries: an emerging and challenging epidemic *J Oncol.* 2010;2010:490631.

10. Porter P. "Westernizing" women's risks: breast cancer in lower-income countries. *New Engj J Med.* 2008;358:213–216.

11. Breast cancer in developing countries *Lancet Oncol.* 2009;374:1077–1085.

12. Aziz Z, Iqbal J, Akram M. Effect of social class disparities on disease stage, quality of treatment and survival outcomes in breast cancer patients from developing countries *Breast J.* 2008;14:372–375.

13. Mittal S. A survey on optimized implementation of deep learning models on the NVIDIA Jetson platform *J Syst Architect.* 2019;97:428–442.

14. Coughlin SS, Ekwueme DU. Breast cancer as a global health concern. *Cancer Epidemiol.* 2009;33:315–318.