# Today's radiologists meet tomorrow's AI: the promises, pitfalls, and unbridled potential

Dianwen Ng[1]^, Hao Du[1]^, Melissa Min-Szu Yao[2,3]^, Russell Oliver Kosik[3], Wing P. Chan[2,3,4#]^, Mengling Feng[1,4#]^

[1]Saw Swee Hock School of Public Health, National University Health System, National University of Singapore, Singapore; [2]Department of Radiology, Wan Fang Hospital, Taipei Medical University, Taipei; [3]Department of Radiology, School of Medicine, College of Medicine, Taipei Medical University, Taipei; [4]Medical Innovation Development Center, Wan Fang Hospital, Taipei Medical University, Taipei

[#]These authors contributed equally to this work.

*Correspondence to:* Wing P. Chan, MD. Department of Radiology, Wan Fang Hospital, Taipei Medical University, 111 Hsing Long Road, Section 3, Taipei. Email: wingchan@tmu.edu.tw; Mengling Feng, PhD. Saw Swee Hock School of Public Health, National University Health System, National University of Singapore, Tahir Foundation Building 12 Science Drive 2, #10-01, Singapore 11754. Email: ephfm@nus.edu.sg.

**Abstract:** Advances in information technology have improved radiologists' abilities to perform an increasing variety of targeted diagnostic exams. However, due to a growing demand for imaging from an aging population, the number of exams could soon exceed the number of radiologists available to read them. However, artificial intelligence has recently resounding success in several case studies involving the interpretation of radiologic exams. As such, the integration of AI with standard diagnostic imaging practices to revolutionize medical care has been proposed, with the ultimate goal being the replacement of human radiologists with AI 'radiologists'. However, the complexity of medical tasks is often underestimated, and many proponents are oblivious to the limitations of AI algorithms. In this paper, we review the hype surrounding AI in medical imaging and the changing opinions over the years, ultimately describing AI's shortcomings. Nonetheless, we believe that AI has the potential to assist radiologists. Therefore, we discuss ways AI can increase a radiologist's efficiency by integrating it into the standard workflow.

**Keywords:** Artificial intelligence; deep learning; diagnostic imaging; radiologists

## Introduction

The role of a radiologist in image-based diagnostics is constantly changing. It has evolved from mundane pattern recognition to the quantitative assessment of imaging features and interventional therapeutic procedures. Advances in information technology have improved radiologists' abilities to perform a variety of targeted diagnostic exams. Furthermore, as populations age across the globe, many countries have implemented a series of new national health initiatives that have increased the volume of screening exams for various diseases. Consequently, the demand for diagnostic imaging has burgeoned, resulting in an excessive workload for radiologists. Concurrently, great strides have been made in the field of artificial intelligence (AI), with AI reportedly surpassing humans' abilities to handle hugely complex tasks. For example, the

---

^ ORCID: Dianwen Ng, 0000-0002-0146-7978; Hao Du, 0000-0002-5083-8122; Melissa Min-Szu Yao, 0000-0002-1295-4771; Wing P. Chan, 0000-0001-8322-6336; Mengling Feng, 0000-0002-5338-6248.

bot AlphaGo beat Ke Jie, the world's best human player of the board game Go, in 2017 (1). Since then, there has been much hype regarding the integration of AI with diagnostic imaging to revolutionize medical care, with the ultimate goal of replacing human radiologists with AI 'radiologists'. Such a proposal sounds attractive, as it could potentially free overworked radiologists to focus on improving patient outcomes through other means. However, such hype should be tempered by skepticism as it is too early to know for certain whether AI will improve patient care.

## Discussion

### Scrutinizing the hype surrounding AI's contributions to diagnostic imaging

Clinical tasks in diagnostic imaging are much more complex than they may seem on the surface. While AI has certainly found success in various intricate tasks like playing Go or understanding humans' natural language, repurposing the algorithms for diagnostic imaging is not straightforward. Most deep learning programs aim to solve a deterministic problem and think in a limited capacity. In medicine, every patient is unique, and the severity of a pathological condition can depend on several factors including age, comorbid diseases, and patient history. AI that only reads medical images lacks the capacity to objectively predict patient outcomes. Furthermore, instruments used to capture medical images produce image noise that can influence what and how the machine learns. This issue has recently been raised by data scientists who study the problem of adversarial attacks (2), that is to say where AI's performance is adversely affected due to small imperceptible perturbations. In addition, the complexity of medical imaging is often underestimated. This can be demonstrated by examining diagnostic mammograms, which are used for the detection of breast cancer (3). Identifying breast cancer on a mammogram is much more than simply identifying the presence of a mass and/or calcifications. Multiple additional factors including morphology, distribution, arrangement, density, and size must all be factored into the assessment. However, typical AI algorithms used in mammography ignore arrangement and location (4), which ultimately affects the performance of the software. Therefore, optimism stemming from the recent successes of AI is excessive, and in designing AI for medical tasks, we should keep in mind the complexity of the task.

AI algorithms do not have the ability to generalize their functions to domains outside of their primary design. Such narrowness means that AI algorithms are only able to make predictions or decisions related to a single task or a limited range of tasks (5). AI software that can detect pathological conditions on chest radiographs cannot detect the same conditions on chest CT scans. Moreover, AI fails to identify outlier cases or cases that differ significantly from the cases used to train the algorithm. For example, an algorithm trained to detect chest infections prior to the current COVID-19 outbreak is of limited utility for the detection of COVID-19, as COVID-19 cases were not used to train the algorithm. Hence, AI can never replace the radiologist's familiarity with the abstract and the anomalous.

In medical imaging, AI works by discovering patterns of data or image features on diagnostic scans. This requires a supervised deep learning method called convolutional neural networks (CNNs) (4). The motivation for the use of CNNs in medical imaging stemmed from their outstanding performance on the ImageNet challenge, where pictures of random objects such as dogs and cars were properly classified with more than 96% accuracy. However, despite the high level of accuracy achieved on the ImageNet challenge, medical imaging tasks pose unique challenges over a different domain. One major challenge lies in the availability of appropriately annotated medical imaging data. Data availability requires patient consent, and confidentiality must be maintained. Researchers must also address the challenges of adhering to the varying standards and regulations imposed by the many institutional review boards. Consequently, without access to real-world data, models will fail to achieve acceptable accuracy. Given these natural constraints of medical imaging, it is likely too early to consider replacing the radiologist with AI.

Having said that, if a large amount of high-quality data could be obtained, the application of AI to medical imaging could be very fruitful. According to a survey published by Biswas, Mainak, *et al.*, the complexity of medical imaging tasks able to be handled by AI has grown rapidly (6). To date, many successful cases of the application of AI to medical imaging have been reported. For example, Esteva *et al.* (7) investigated the ability of a deep CNN model trained with a large dermoscopy dataset to discriminate between the most common skin cancers. The model matched the performance of 21 board-certified dermatologists in evaluating malignancy of the skin. The authors suggested that mobile devices such as smartphones could be deployed with similar algorithms, potentially permitting low-cost universal access to vital diagnostic care anywhere in the world. In this

area, AI can greatly reduce healthcare costs and provide an early warning to approximately 9,500 people in the US who are at risk of skin cancer every day. Such work offers clear evidence of AI's efficacy and the potentially boundless possibilities for its application to medical imaging.

Nonetheless, the hype around AI has waned somewhat. Among the success stories, there have been some examples in recent years where AI has failed to meet expectations. For example, from 2018 to 2019, Google conducted the first clinical trial of an AI algorithm used to assess retinal images (8). The AI software developed by Google Health can identify signs of diabetic retinopathy from eye scans with an accuracy rate of more than 90% (the team calls it "human expert level"), and in principle, it can provide a patient with results in less than 10 minutes. However, the clinical trial revealed a number of limitations and concerns. The data used to train the AI algorithm were of high quality. Low-quality images were rejected. Because nurses scan dozens of patients every hour and often do so in low lighting, more than one-fifth of the images were deemed low-quality by the algorithm and subsequently rejected (9). In addition, poor internet connections in some clinics also caused problems. The retinal images had to be uploaded to a cloud server for the AI algorithm to perform its analysis. The slow upload speed resulted in significant, which were far from the originally expected instant results. Thus the hype around AI in medical imaging has cooled some. AI certainly still has great potential, but a number of limitations and challenges still must be solved (10).

### Radiologists' new responsibilities in the AI-assisted healthcare ecosystem

With the communal efforts driving research into the application of AI to medical imaging (11), we envision that the technology will improve the healthcare system in many ways. However, AI cannot do it alone. Rather, every radiologist must be involved in the process. That is to say, radiologists make decisions that can have great health and financial consequences. One of the biggest limitations of AI is its lack of human intelligence skills (intuition, cognition, and reasoning). Machines may make baffling recommendations that could prove harmful if unquestioningly followed. For example, the chest X-rays of the 5,302 pneumothorax patients in the NIH database often contain foreign objects such as chest tubes. During training, the machine sometimes associates the presence of a chest tube with the presence of a pneumothorax. Deep learning-based AI models are often considered black box solutions that can be hard for humans to understand and trust. We cannot determine the extent to which AI learns features correctly, and catastrophe can potentially occur if learning is fundamentally wrong. Therefore, the radiologist must always act as the final judge in hospitals that employ AI. The doctor should check for errors in cases that are difficult to interpret or misleading.

However, radiologists can exploit the accuracy of AI models as a second opinion to boost their confidence. This is advantageous to every radiologist and could be even more useful to junior doctors, given that models learn directly from senior consultants who have years of experience. To combat "narrow" AI (5), we can train multiple AI algorithms that each master different parts of the body. AI must adopt technical knowledge in 3 main spheres: classifying cases, generating diagnostic reports, and providing treatment recommendations. Subsequently, junior doctors can utilize AI algorithms at work or as a virtual learning platform where they can engage AI at their convenience and receive input regarding interesting cases. Such an environment of knowledge-sharing and reinforcement will no doubt facilitate young radiologists' learning, and their grasp of more complex medical imaging techniques and principles can be expected to increase.

Lastly, as we deploy AI-integrated systems to assist the radiologist, we must not forget to perform timely and routine audits to check and retrain the models so that they remain up to date. The radiologist is tasked with regularly generating labels for new and, most importantly, rare and difficult, cases. This will establish a perpetual feedback loop between radiologists who receive assistance from machines and machines that receive new data, thus improving their ability to interpret medical images (*Figure 1*). However, current FDA regulations do not allow temporal updates to AI models (12). At this juncture, models require new approval following every update. While we must be prepared to discontinue the AI software whenever integrity checks of its performance fail, we nonetheless hope for some loosening of the FDA policies regarding temporal updates so as to ensure that algorithms are functioning at their best.

### The future

In this age of technology, it is predicted that AI applications will help manage and save resources that can then be channeled to more efficient uses. AI might be a way to
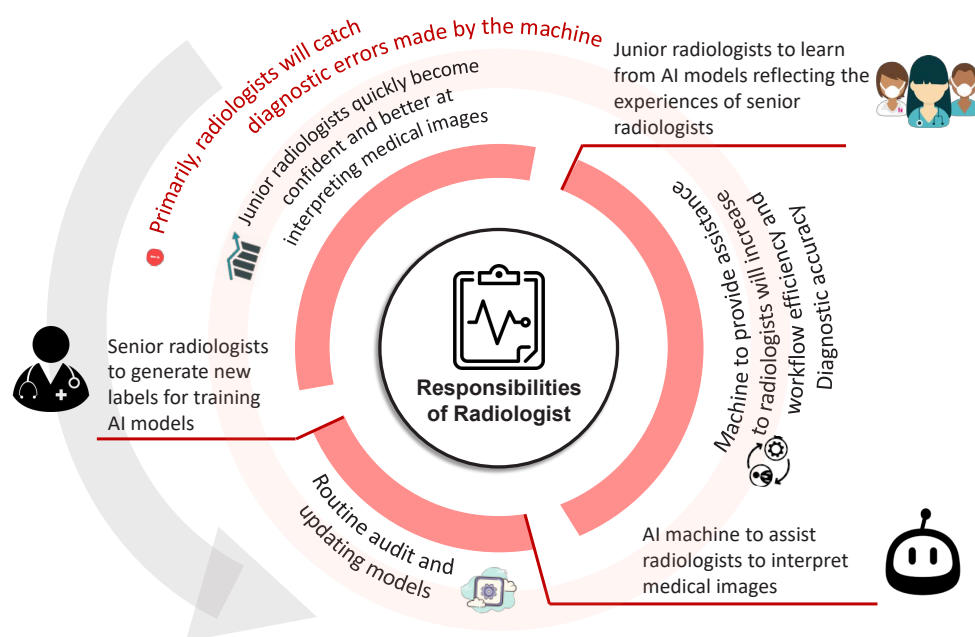
**Figure 1** Schematic showing the role of the future radiologist in diagnosis and management with the assistance of artificial intelligence as proposed in this paper.

curb the rising costs of healthcare and augment the quality of care. Radiologists can then spend more time on the diagnosis and management of difficult and rare cases, where AI is expected to fail. In addition, our radiologists can pay more attention to research that further advances medical imaging technologies. The radiologists of tomorrow should not resist integrating AI into the healthcare ecosystem. While it is still unclear how exactly AI will impact their daily work, radiologists should become more skilled at interacting with artificially intelligent machines in a way that reaps their full benefits.

## Acknowledgments

## Footnote

of interest to declare.

*Ethical Statement:* This article does not contain any studies with human participants or animals performed by any of the authors.

## References

1. Tam PW. Daily report: AlphaGo shows how far artificial intelligence has come. New York Times. May 23, 2017. Available online: https://www.nytimes.com/2017/05/23/technology/alphago-shows-how-far-artificial-intelligence-has-come.html. Accessed September 28, 2019
2. Ma X, Niu Y, Gu L, Wang Y, Zhao Y, Bailey J, Lu F. Understanding adversarial attacks on deep learning based

medical image analysis systems. Pattern Recognition. 2020 May 1:107332. arXiv:1907.10456.

3. The Radiology Assistant : Bi-RADS for Mammography and Ultrasound 2013. Available online: https://radiologyassistant.nl/breast/bi-rads/bi-rads-for-mammography-and-ultrasound-2013. Accessed January 22, 2021

4. LeCun Y, Bengio Y. Convolutional networks for images, speech, and time series. The handbook of brain theory and neural networks 1998:255-8.

5. Ahuja AS. The impact of artificial intelligence in medicine on the future role of the physician. PeerJ 2019;7:e7702.

6. Biswas M, Kuppili V, Saba L, Edla DR, Suri HS, Cuadrado-Godia E, Laird JR, Marinhoe RT, Sanches JM, Nicolaides A, Suri JS. State-of-the-art review on deep learning in medical imaging. Front Biosci (Landmark Ed) 2019;24:392-426.

7. 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;542:115-8.

8. Beede, Emma and Baylor, Elizabeth and Hersch, Fred and Iurchenko, Anna and Wilcox, Lauren and Ruamviboonsuk, Paisan and Vardoulakis, Laura M. "A Human-Centered Evaluation of a Deep Learning System Deployed in Clinics for the Detection of Diabetic Retinopathy." Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems. April 2020 Pages 1–12. Available online: https://doi.org/10.1145/3313831.3376718

9. Beede, Emma. Healthcare AI systems that put people at the center. Available online: https://www.blog.google/technology/health/healthcare-ai-systems-put-people-center. Accessed January 22, 2021

10. Heaven WD. Google's medical AI was super accurate in a lab. Real life was a different story. MIT Technology Review. Available online: https://www.technologyreview.com/2020/04/27/1000658/google-medical-ai-accurate-lab-real-life-clinic-covid-diabetes-retina-disease. Accessed January 22, 2021

11. Sahiner B, Pezeshk A, Hadjiiski LM, Wang X, Drukker K, Cha KH, Summers RM, Giger ML. Deep learning in medical imaging and radiation therapy. Med Phys 2019;46:e1-36.

12. Center for Devices and Radiological Health. Artificial Intelligence and Machine Learning in Software. U.S. Food and Drug Administration. FDA; Available online: https://www.fda.gov/medical-devices/software-medical-device-samd/artificial-intelligence-and-machine-learning-software-medical-device. Accessed January 22, 2021