From identifying dog breeds to diagnosing diabetic retinopathy
How artificial intelligence can revolutionize medical diagnosis
Shafaz Veettil, Logan Van Nynatten

ABSTRACT
Experts predict advances in artificial intelligence (AI)—the ability of a machine to mimic human cognition—will spark the fourth industrial revolution, but the dawn of a new age in diagnostic medicine may already be on the horizon. Google and others are leveraging deep learning, a subset of AI that aims to imitate the neurological processing of the human brain, to screen for diseases such as diabetic retinopathy, cardiovascular disease, brain tumours, skin cancers, and stroke, at unprecedented levels of sensitivity and specificity. Limitations include an inability to wholly substitute for human empathy and touch, risk for adversarial training, and concerns about interpretability. If hurdles can be properly dealt with, the coalescence of big data and AI research will change how medicine is provided, practiced, and accessed.

INTRODUCTION
Artificial intelligence (AI)—the ability of a machine to mimic human cognition—is poised to have a transforming impact upon society. Today, it underpins our search engines and speech recognition software; one day, it will power our self-driving cars and herald the fourth industrial revolution. Current advances in AI technology are predicated on machine learning and, more specifically, deep learning. Machine learning is an AI technique that involves algorithms that learn from data to make informed decisions. Deep learning is an evolution of this approach and involves the use of neural networks—layered algorithms inspired by the hierarchal nature of the human brain—to facilitate higher-level learning and conclusions. When classifying dog images, machine learning techniques would be able to identify dog breeds from standardized images, while deep learning machines would exploit multilevel abstract processing to recognize subtle differences in dog breeds from images taken in a variety of settings.

Increasingly, the use of AI is also augmenting our ability to diagnose disease. With the advent of deep learning, for the first time machines can meet or outperform the diagnostic standards-of-care. The appliance of AI will change how physicians diagnose disease; examples and what that means for the doctor-patient relationship, privacy, and the medical profession will be examined below.

APPLICATIONS
Medical researchers, often in partnership with private companies, are currently experimenting with AI technology in the context of diagnosing important diseases. Google’s work with diabetic retinopathy is one such example. Diabetic retinopathy is a leading cause of vision loss in Americans and well-suited for screening; however, the large number of retinal images that would need to be evaluated has prevented screening for this disease from being routinely practiced. In response, a team from Google has employed an automated deep learning algorithm to make screening more practical. Trained from 128,175 images and validated by a panel of US board-certified ophthalmologists, the algorithm achieved 97.5% sensitivity and 93.4% specificity, levels comparable to ophthalmologists and an improvement over those attained by general physicians and optometrists. Another group of Google employees leveraged deep learning models to precisely extract key cardiovascular risk factors (age, gender, smoking status, systolic blood pressure, and major adverse cardiac events) simply from retinal fundus photographs; this presents a quick, inexpensive, and non-invasive potential approach to risk-stratifying patients for cardiovascular disease. Microsoft and IBM are also involved in this sphere; Microsoft’s InnerEye project is working to diagnose and characterize brain tumours from CT scans using deep learning techniques while IBM is employing Watson to achieve personalized medicine in oncology.

The versatility of AI may mean a future where it is integrated into all aspects of healthcare: AI is already in use for early stroke detection via a device that analyzes for abnormal patient movement patterns, for stroke diagnosis from neuroimages, and in the emergency setting to differentiate stroke from stroke mimics based on patient presentation, risk factors, and other clinical aspects. The potential scalability of AI has led Esteva et al.—who developed AI software that can classify skin lesions at levels on par with dermatologists—to envision their technology reaching billions of people through the universality of smartphones.

LIMITATIONS
As the marriage of medicine and AI becomes more profound, questions pertaining to the shortcomings of the technology will need to be answered or at least appreciated. Although the potential of AI may seem all-encompassing, boundaries in the scope of the technology exist. In its present state, AI cannot engage in meaningful conversations, forge trust, and express empathy. Without these skills—drivers of patient satisfaction and health outcomes, and components of what we call the ‘art of medicine’—the provision of humanity to patients will remain in the realm of the physician. For similar reasons, the healthcare system will continue to rely on physicians to fulfill areas in which histories and physical
exams are needed for diagnosis—such as in rheumatology, where the importance of patient history and physical examination can significantly outweigh that of laboratory and ancillary tests, and in neurology, where complex interactions are required—and to bear the responsibility of patient follow-up. In addition, it is important to highlight that the medical knowledge that an AI uses to make or aid in a diagnosis stems from that supplied by medical experts; these medical specialists will be needed to monitor a system’s validity and ensure that it is updated with the latest knowledge. Finally, despite the tremendous progress already made in the field, sensitivity and specificity figures will need to be improved upon to reduce the risk of misdiagnoses and ensure acceptable standards of care are met from clinical, legal, and the patient’s perspectives before AI is ready to replace trained physicians.

The question of validity is especially important when one considers the issues of bias and interpretability. AI—like a human—is vulnerable to adversarial training; if the dataset an AI is trained on is biased, then the output diagnosis produced by the AI will be biased. This potential for biased algorithms is illustrated by the findings of Wiley et al that show warfarin-dosing algorithms underperform in African Americans because the clinical trials the algorithms were derived from lacked information on the pharmacogenomic differences of this population. Furthermore, a lack of data on rare diseases or disease variants would cause AI systems to overlook such cases. Research subjects are disproportionately ill, male, and Caucasian, and for an AI system to be applied universally, designers of the system will need to ensure the diversity and generalizability of the training dataset, or at least specify the system’s scope. Such patient training datasets, in light of the growing number of collaborations between public and private industries in AI research and development, will also need to be protected and deidentified to safeguard against privacy breaches.

The inherent risk of bias and adversarial training also warrants a way for clinicians to interpret the recommendations produced by the system. Many present models of AI are “black box”, with incredibly complex algorithms outputting a diagnosis or prognosis from the input. Without the capability to make sense of or truly interpret how input progresses to output, physicians and patients must place their full trust into machines that may inherit our biases. Such a reality is uncomfortable to some, and places the question of who—the physician, the hospital, or the designers of the AI system—would be most responsible in a malpractice suit in such a scenario. Alternative models involving supervised machine learning can offer greater algorithmic transparency and insight than deep learning models but require more resources and collaboration between physicians and computer scientists.

**IMPLICATIONS**

AI has profound implications for physicians, patients, and policymakers. Parties will need to grapple with the concerns outlined above—privacy hurdles, bias, trust and interpretability—and must work together to introduce legal and professional regulations that protect against any currently unforeseen deterioration of healthcare standards and safeguard against technological abuse. Work will also need to be done to optimize the incorporation of AI into healthcare. According to Wong and Bressler, practical challenges include identifying if patients are best served by AI technology that acts at the point of care (e.g. at an ophthalmologist’s clinic) or at centralized reading centers. To encourage adoption, AI software should be inexpensive, quick, flexible, user-friendly, and able to articulate the algorithmic calculus behind the diagnosis-making process.

The automating power of AI is seen as a disruptive force in human employment, and the field of medicine will likely not be exempt from its influence. Though the specialties of radiology and pathology seem most at risk of replacement from AI, according to Jha and Topol, the future of these professions may be rosier than many expect. Jha and Topol argue that AI will be a boon; though the adoption of AI will amplify the already large volumes of imaging data that need to be processed, it will also free radiologists and pathologists from the rote analysis of such data and allow them to act in a combined role known as an “information specialist.” Someone in this capacity would manage and integrate information produced by AI and use it to guide clinicians in the context of the patient. In the same way that today’s pathologists have automated tasks such as cell counting, tomorrow’s information specialists may rely on AI to augment their practice and allow them to operate on much larger economies-of-scale than ever before.

As mentioned earlier, as long as the physician-patient relationship, patient history taking, and physical examinations are of value, elements of a clinician’s job will remain secure. Moreover, in many ways, AI may make the life of a clinician easier. According to Frost and Sullivan, AI is capable of improving health outcomes by 40% and reducing expenses by half. The benefits of AI may be particularly advantageous in underserved and low-resource areas, where there are inadequate number of medical specialists, as primary care physicians will be empowered with the tools necessary to make diagnostic decisions without having to consult specialists directly. Another consequence of this empowerment would be the reduction of wait times to see these specialists as only a curated set of complex patients will require the attention of a specialist. Conversely, any cost- and infrastructure-issues that delay adoption of AI technology in low-resourced settings would contribute to disproportionalities in national and global health.

**CONCLUSION**

In order for the technology to be optimally implemented, stakeholders will need to be educated on the strengths of the technology, such as analytics and image recognition, as well as its deficiencies, including the human elements of medicine, the lack of immunity to bias, and the “black box” nature of its algorithms. There is still a long way to go before many of the weaknesses can be rectified—and a gaping need for regulations to develop in parallel with AI advances—but the future is bright; the capability of AI to be trained to identify anything from dog breeds to diabetic retinopathy will make it a valuable and paradigm-shifting fixture in diagnostic medicine.
REFERENCES


