

Artificial Intelligence and the Pathologist

Future Frenemies?

Gaurav Sharma, MD; Alexis Carter, MD

The manuscript titled “AlphaGo, deep learning, and the future of the human microscopist” in this month’s issue of the *Archives of Pathology & Laboratory Medicine*¹ describes the triumph of Google’s (Mountain View, California) artificial intelligence (AI) program, AlphaGo, which beat the 18-time world champion of Go, an ancient Chinese board game far more complex than chess.

The authors have hypothesized that the development of intuition and creativity combined with the raw computing of AI heralds an age where well-designed and well-executed AI algorithms can solve complex medical problems, including the interpretation of diagnostic images, thereby replacing the microscopist. Of note, in a prior work, the microscope was predicted to have a 75% chance of remaining in use for another 144 years.²

To support their hypothesis, the authors presented recent studies that compared the performance of nontraditional interpreters to those of experienced pathologists, in making accurate diagnoses (note: 1 author disclosed a significant financial interest in an AI company). One study examined the potential of using pigeons (yes, pigeons) for medical image studies,³ wherein the pigeons engaged in a matching game of completely benign and unambiguously malignant breast histology images. Pigeons correctly classified images as benign or malignant 85% of the time. A separate image algorithm study was erroneously reported to differentiate between small cell and non-small cell lung carcinoma with the accuracy of expert pulmonary pathologists, but instead, multiple computational algorithms were used to subtype known non-small cell lung carcinomas and gliomas in separate experiments.⁴ The accuracy rate of each algorithm approached 70% to 85%. We believe that this level of diagnostic accuracy in settings that lack complexity is an extremely poor replica of a human pathologist’s diagnostic capabilities.

So, will the data-digesting and 24×7 learning AI be capable of looking at an image and able to render a pathologic diagnosis? Before attempting to answer this, we caution against the difficulties of predicting the future. Much of our existence still rests on innovations that have remained unchanged because of their inherent simplicity, applicability, and trueness to purpose (eg, the wheel), proving the point that something new (and different) is not always something better. On the other hand, several established and incumbent technologies were quickly (albeit incompletely) eclipsed, often within a decade, by a challenger that was faster, more convenient, cheaper, or better for the need (eg, postal mail being replaced by electronic mail). In the latter context, we note that information technology and AI are clearly better at repetitive detailed tasks that require accuracy and speed than are humans, who often find such tasks mind-numbing, and consequently are error-prone.

Also see p. 619 and p. 624.

Interestingly, IBM’s Watson (Armonk, New York) and Google’s AlphaGo attempt to recreate (and outpace) human thought process steps: observation, initial interpretation, hypothesis generation, evaluation, and decision-making. In fact, IBM’s Watson is most famous for its defeat of the 2 highest-winning “Jeopardy!” champions of all time, and Watson’s purpose was always health care.⁵ One example of this is Watson Genomics, an assistive AI to improve next-generation sequencing (NGS) interpretations.⁶

Perhaps, rather than question diagnostic equivalence, the better question is will AI eventually be capable of replicating (and thus replacing) the professional role of the pathologist? To this question, we believe that the answer is still a “no”—because this question is an erroneous comparison between 2 very dissimilar activities—high-level cognition (a human forte) versus high-level computation (an AI forte, at least for now). We believe that a pathologic diagnosis is often a well-thought-out cognitive opinion, benefiting from our training and experience, and subject to our heuristics and biases. We submit that as pathologists, our professional value comes from our ability to give the most appropriate (even if not the perfect) opinion that amalgamates the available clinical contexts and the clinical questions—2 imperfect, variable, and nonuniform inputs. Further, as humans, we can overcome the nuances of human communication to constantly recalibrate our diagnosis based on small but significant nuggets of clinical and patient-specific informa-

Accepted for publication December 20, 2016.

From the Department of Pathology & Laboratory Medicine, Henry Ford Health System, Detroit, Michigan (Dr Sharma); and the Department of Pathology and Laboratory Medicine, Children’s Healthcare of Atlanta, Atlanta, Georgia (Dr Carter).

The authors have no relevant financial interest in the products or companies described in this article.

doi: 10.5858/arpa.2016-0593-ED

Reprints: Gaurav Sharma, MD, Department of Pathology & Laboratory Medicine, Henry Ford Health System Room W631A, K6 Pathology, 2799 West Grand Blvd, Detroit, MI 48202 (email: Gsharma2@hfhs.org).

tion assimilated from different data rooted in the human language, such as physician's notes, pathology reports, email, and verbal discourse with our clinical peers.

We do feel that eventually, our cognitive lead will narrow as AI products like IBM Watson demonstrate valid cognition of patient and population information. In the interim, because the language-based foundation of medicine is unlikely to disappear anytime soon, it would be more reasonable to see NGS, digital pathology, whole-slide imaging, and AI as synergistic technologies to human cognition. We note that the question of human versus computer has now been refined as, human versus human with the computer. The verdict is clearly in latter's favor, as described in Clinical Informatics' Fundamental Theorem: "A person working in partnership with an information resource is better than that same person unassisted."⁷

The novelist Gertrude Stein observed: "This is the lesson that history teaches: repetition."⁸ Ergo, before we feel that AI will take over pathology, we should take a lesson from history: the evolution of clinical decision support (CDS) systems in clinical medicine. In the 1970s, the challenges of expanding medical knowledge, shortage of specialists, and improving information technology capabilities prompted the development of algorithms that could serve as diagnostic "experts." Examples include the Leeds Abdominal System, MYCIN, and INTERNIST-1.⁹⁻¹¹ Some of these systems' diagnostic accuracy eclipsed that of the clinicians (92% versus 65%–80% accuracy, respectively, for the Leeds system). Claims that the computers were poised to take over the medical profession were commemorated in the series "Star Trek: Voyager" wherein the medical doctor was an "Emergency Medical Hologram."¹² However, by the late 1980s, the clinical community had realized that the diagnostic process was too complicated and diverse to be trusted to hard-wired algorithms. Not surprisingly, nowadays the focus of CDS systems has evolved away to being supportive tools incorporated within the electronic health record (EHR), accelerated by requirements under meaningful use. The vast majority of hospitals now operate EHR systems with a repertoire of CDS assistive (rather than diagnostic) features, such as alerts for adverse drug events and immunization reminders.¹³

The promulgation of AI technology will have to hurdle several significant barriers not previously mentioned by the authors. Many assistive technologies are not available to health care providers because instruments, laboratory information systems, and/or EHR systems have failed to incorporate/interface them. For example, many high-tech molecular devices cannot read bar codes or interface with a

laboratory information system. Even after decades of the existence of interoperability standards, such as HL7, health systems continue to struggle with simple access to patient information. Similar to NGS, digital pathology and whole-slide imaging, the financial benefit and scalability for AI remain undefined and may, ironically, depend on human ability to reach agreements to make these technologies agnostic to proprietary platforms.

We conclude by mentioning Dr Fill, an AI entity potentially capable of reproducing the most complex of cognitive products—a human's sense of humor. Dr Fill's creator, Matthew Ginsberg, said, "... computers have natural domains of competence that are very different from ours. And it's good that we're different because it means we're not natural competitors, we're natural cooperators."¹⁴ Therefore, we believe that the metrologic forecast favors a sunny era of AI assistance in pathology rather than dark clouds of AI competition replacing pathologists—stay tuned for updates!

References

1. Granter SR, Beck AH, Papke Jr DJ. AlphaGo, deep learning and the future of the human microscopist. *Arch Pathol Lab Med*. 2017;141(5):619–621.
2. Granter SR. Reports of the death of the microscope have been greatly exaggerated. *Arch Pathol Lab Med*. 2016;140(8):744–745.
3. Levenson RM, Krupinski EA, Navarro VM, Wasserman EA. Pigeons (*Columba livia*) as trainable observers of pathology and radiology breast cancer images. *PLoS One*. 2015;10(11):e0141357.
4. Hou L, Samaras D, Kurc TM, Gao Y, Davis JE, Saltz JH. Patch-based convolutional neural network for whole slide tissue image classification. *Proc IEEE Comput Soc Conf Comput Vis Pattern Recognit*. 2016;2016:2424–2433.
5. IBM Watson Health. Watson Health: welcome to the cognitive area of health. <http://www.ibm.com/watson/health/>. Accessed November 28, 2016.
6. IBM Watson Health. IBM Watson for Genomics. <http://www.ibm.com/watson/health/oncology/genomics/>. Accessed November 22, 2016.
7. Friedman CP. A "fundamental theorem" of biomedical informatics. *J Am Med Inform Assoc*. 2009;16(2):169–170.
8. This is the lesson that history teaches: repetition. AZQuotes. <http://www.azquotes.com/quote/282254>. Accessed December 12, 2016.
9. Horrocks JC, Devroede G, de Dombal FT. Computer-aided diagnosis of gastroenterologic diseases in Sherbrooke: preliminary report. *Can J Surg*. 1976;19(2):160–164.
10. Miller RA, Pople HEJ, Myers JD. Internist-1, an experimental computer-based diagnostic consultant for general internal medicine. *N Engl J Med*. 1982;307(8):468–476.
11. Yu VL, Fagan LM, Wraith SM, et al. Antimicrobial selection by a computer: a blinded evaluation by infectious diseases experts. *JAMA*. 1979;242(12):1279–1282.
12. The Doctor (Star Trek: Voyager). Wikipedia. [https://en.wikipedia.org/wiki/The_Doctor_\(Star_Trek:_Voyager\)#Emergency_Medical_Hologram](https://en.wikipedia.org/wiki/The_Doctor_(Star_Trek:_Voyager)#Emergency_Medical_Hologram). Accessed October 20, 2016.
13. Health Information Technology for Economic and Clinical Health (HITECH) Act, Title XIII of Division A and Title IV of Division B of the American Recovery and Reinvestment Act of 2009 (ARRA), Pub L No. 111-5, 123 Stat 226 (Feb. 17, 2009), codified at 42 USC §§300jj et seq; §§17901 et seq. <https://www.gpo.gov/fdsys/pkg/BILLS-111hr1enr/pdi/BILLS-111hr1enr.pdf>. Accessed December 12, 2016.
14. Bennet D. Artificial intelligence? I'll say: why computers can't joke. <http://www.bloomberg.com/news/articles/2012-03-23/artificial-intelligence-ill-say-dot-why-computers-cant-joke>. Accessed November 28, 2016.