Published online: 2019-04-25

© 2019 IMIA and Georg Thieme Verlag KG

Artificial Intelligence in Medicine: Weighing the Accomplishments, Hype, and Promise

Edward H. Shortliffe
Department of Biomedical Informatics, Columbia University
College of Health Solutions, Arizona State University
Department of Health Policy and Research, Weill Cornell Medical College

Summary

Introduction: Artificial Intelligence in Medicine (AIM) research is now 50 years old, having made great progress that has tracked the corresponding evolution of computer science, hardware technology, communications, and biomedicine. Characterized as being in its "adolescence" at an international meeting in 1991, and as "coming of age" at another meeting in 2007, the AIM field is now more visible and influential than ever before, paralleling the enthusiasm and accomplishments of artificial intelligence (AI) more generally.

Objectives: This article summarizes some of that AIM history, providing an update on the status of the field as it enters its second half-century. It acknowledges the failure of AI, including AIM, to live up to early predictions of its likely capabilities and impact.

Methods: The paper reviews and assesses the early history of the AIM field, referring to the conclusions of papers based on the meetings in 1991 and 2007, and analyzing the subsequent evolution of AIM.

Conclusion: We must be cautious in assessing the speed at which further progress will be made, despite today's wild predictions in the press and large investments by industry, including in health care. The inherent complexity of medicine and of clinical care necessitates that we address issues of usability, workflow, transparency, safety, and formal clinical trials. These requirements contribute to an ongoing research agenda that means academic AIM research will continue to be vibrant while having new opportunities for more interactions with industry.

Keywords

Artificial intelligence, machine learning, decision-support systems, history of artificial intelligence in medicine

Yearb Med Inform 2019:257-62
http://dx.doi.org/10.1055/s-0039-1677891

1 Introduction

The headlines abound: "AI in healthcare market to grow to US$36.1B by 2025" [1], "6 ways AI is changing healthcare" [2], "Artificial Intelligence Transforms the Future of Medicine" [3], "AI Shows Promise Assisting Physicians" [4]. A drive up California's Bayshore Freeway to San Francisco currently reveals multiple billboards for Silicon Valley commuters, many touting AI applications and tools from major corporations and startups (many of which are health-related), while others indicate the high-paying job opportunities for people with training in AI and data analytics. We talk to our smart phones and to virtual assistants in our homes and cars, while radiologists dictate into computer systems that generate text reports automatically. Futurists and medical pundits envision a new world of "high performance medicine", resulting from the convergence of human and artificial intelligence [1].

For those of us who have been engaged with research on artificial intelligence in medicine and health care for decades, the current visibility and enthusiasm regarding our field is both refreshing and frightening. It is refreshing for those who worked hard to create systems that would demonstrate the promise of the field – often with pushback from the medical community, who could not initially envision the use of computers in health care that has gradually become commonplace. But it is frightening for those of us who remember the overselling of AI in general, including medical AI, in the early 1970s, late 1980s, and early 1990s. The hype in the press in the 1980s led first to societal curiosity, then to unfounded predictions by observers, and ultimately to the phenomenon of "AI winter" that followed when predictions proved to be overly optimistic. With a drop in research funding for anything related explicitly to AI, there was a tendency for those working in the field to select new terms for the science they were pursuing. Knowledge representation researchers developed *ontologies* and the machine learning community gathered at meetings called *knowledge discovery in databases* (KDD).

At the 1991 *AI in Medicine Europe* (AIME) meeting, the organizers asked me to present an assessment of the current state of AI in Medicine (often shortened to AIM). I characterized the field as being in its adolescence, with some solid accomplishments but many challenges, coupled with misunderstandings about the nature and goals of the field [2]. Subsequently, at the 2007 AIME meeting, several of us convened an invitational panel to consider progress in the field in the intervening years, later writing up our comments to convey our sense that AIM was coming of age [3]. Although we acknowledged the unfortunate impact of the overselling of AIM in earlier years, we made

[1] https://www.beckershospitalreview.com/artificial-intelligence/ai-in-healthcare-market-to-grow-to-36-1b-by-2025.html (Accessed Feb 10 2019)

[2] https://www.beckershospitalreview.com/artificial-intelligence/6-ways-ai-is-changing-healthcare.html (Accessed Feb 10 2019)

[3] https://news.aamc.org/research/article/artificial-intelligence-transforms-future-medicine/ (Accessed Feb 10 2019)

[4] https://www.nytimes.com/2019/02/11/health/artificial-intelligence-medical-diagnosis.html (Accessed Feb 12 2019)

the case that the field was actually healthy, growing, and poised to have greater visibility and impact. We pointed to the importance of the increasingly available computational power, ubiquity of computing, networking and interconnectivity, and an environment for health professionals and the public that was more accepting of the notion that computers could play an increasingly valuable support role in health and health care.

In this article, I will offer a Part 3 to this series of papers, observing how another decade has affected our field, addressing the challenges in promoting what is possible while attempting to rein in unrealistic expectations, and anticipating what lies ahead, both for AIM and for the biomedical informatics discipline to which it contributes. I begin with a brief recapitulation of the history of AIM, noting that we provided more detail in those two earlier papers [2, 3]. I also identify key issues that characterize AIM research as we enter the next decade in the evolution of medical AI.

2 The Beginnings of AI and its Relevance to Medicine

The term "artificial intelligence" was coined by John McCarthy in a proposal for a computer science conference that was held at Dartmouth College in 1956[5]. McCarthy was then at the Massachusetts Institute of Technology (MIT) but went on to spend most of his career on the computer science faculty at Stanford University. A web site maintained by the Association for the Advancement of Artificial Intelligence (AAAI) nicely summarizes chronologically the subsequent history of the AI field[6]. In the years that followed that conference, several universities developed research programs in AI, often with support from the Defense Advanced Research Projects Agency (DARPA) in the US Department of Defense. Early centers of excellence emerged at MIT, Stanford, and Carnegie Mellon University, and there were important contributions underway at several other universities as well. The phrase "machine intelligence" was used in the UK, where there were also important research activities. Although much of the early work focused on generalized methods for simulating or reproducing human-like intelligence, by the late 1960s major efforts began to examine how AI methods could be used to tackle specific problem areas in the natural sciences. At Stanford, the DENDRAL Project of geneticist Joshua Lederberg and computer scientist Edward Feigenbaum, in collaboration with organic chemist Carl Djerassi and philosopher-of-science Bruce Buchanan, used production rules to encode the knowledge of expert chemists. They also applied clever algorithms that inferred likely chemical structures of organic compounds from mass-spectral data [4].

Lederberg subsequently led the effort to obtain funding from the National Institutes of Health (NIH) for a shared computer system at Stanford that was to be connected to the still-nascent ARPAnet (predecessor to today's Internet) and would provide a resource for a community of individuals undertaking research in the application of AI in the biomedical and clinical sciences. The resulting SUMEX-AIM computer, and a second networked shared resource at Rutgers University, coordinated to offer services to AIM researchers, initially within the US and later internationally. Regular national AIM workshops facilitated collaborations and the formation of a community that came to know each other's work well. Much of the AIM community's work was summarized in a 1978 special issue of the journal *Artificial Intelligence* [5], in a 1984 book on major early research projects [6], and in a 1980 NIH publication describing federal investments in the area [7]. My own dissertation work on the MYCIN System [8, 9] would have been impossible without Lederberg's leadership [10], Buchanan's guidance, and the creation of the SUMEX-AIM resource on which I carried out my research.

Medical AI projects were responsible in part for the explosion of public interest in expert systems by the early 1980s. New computing machines were developed to support such work, largely built to optimize the LISP programming language. Major companies formed AI research groups and built academic collaborations to explore how the new technologies might stimulate innovation in their businesses. National and international news outlets ran front-page stories highlighting AI and expert systems, often predicting rapid changes that would revolutionize society, with health care often a central focus of the articles. As the decade progressed, however, it became clear that the aspirations and predictions were not yet achieved, with progress that was slow and projects that were constrained by the complexity of what was being undertaken. This disillusionment led to articles about "the failure of AI and expert systems" (AI winter), tainting the reputation of the field even as slow and steady progress continued to be made[7].

By the time of the previously mentioned AIME meeting in Maastricht (1991), the global AIM research community was active and vibrant, but somewhat sidelined and reluctant to characterize what they were doing as AI or expert systems research. Introspection on an international AIM list server led to my summary of, and responses to, key questions being asked in the community:

- Have we been so focused on decision-making performance that we have failed to address user needs? *[Response: The best work had not failed to consider these issues.]*
- Can we identify clear contributions of AIM to the general discipline of AI, to psychology, to clinical medicine? *[Response: Yes, for the first two, but impact on clinical medicine was still to be achieved.]*
- Should we view AIM as part of information systems, computer science, AI itself, engineering, or biomedicine? *[Response: None of the above. It is a key component of biomedical informatics, which is itself a separate hybrid discipline.]*
- Are we training AIM scientists well? *[Response: Only in a few places at that time.]*

[5] https://www.forbes.com/sites/gilpress/2016/08/28/artificial-intelligence-defined-as-a-new-research-discipline-this-week-in-tech-history/#bf5ddbf6dd15 (Accessed Feb 10 2019)

[6] https://aitopics.org/misc/brief-history (Accessed Feb 10 2019)

[7] https://www.basistech.com/honest-ai/the-ai-hype-machine/ (Accessed Feb 12 2019)

- Is there inbreeding or a limitation of perspective? *[Response: Should not be a problem as the field grows and prospers, which has subsequently been borne out.]*
- Why does there appear to be a lack of studies of AIM systems in routine use? *[Response: The problems are difficult, systems can adversely affect workflow, and it is especially hard to introduce systems smoothly in the absence of standards that support integration.]*
- Given the potential importance of the field, why is it poorly funded by research agencies? *[Response: Getting funding can always be challenging, and AI had come into some disfavor, but excellent work was still being supported.]*
- Why are our systems exceedingly difficult to move from one site to another? *[Response: Many reasons, including lack of sensitivity to cultural and process differences across institutions, absence of standard terminologies or connectivity standards, and "not invented here" biases.]*

These questions, and my more detailed answers to them, became the focus of my 1991 AIME presentation and of the subsequent published version to which readers may refer for a then-current (and generally upbeat) analysis [2].

The next 15 years saw remarkable changes both in technology and in the capabilities of AIM systems. Machine learning had become a significant focus of AIM research [11, 12], and the importance of integration with clinical systems (facilitated by the gradual development of pertinent standards) had become well recognized. The biological sciences had embraced AI notions, spurred on by the size of their datasets and the success of the Human Genome Project in the 1990s. Meanwhile, wireless networking, smart phones, and social media had all contributed to a societal acceptance of computing that had been less prevalent even a decade earlier. When we gathered for the previously mentioned AIME session in Amsterdam (2007), organizers asked the five panelists to address how the field had advanced since 1991. They also wanted to know the extent to which it was influencing biomedicine or clinical fields, and how well it was being supported by funding agencies, by academic or research entities, and by colleagues in computer science or biomedical fields. Those perspectives are summarized in what I have called Part 2 in this series of papers [3]. In my own comments I stressed that AIM could not be set off from the rest of biomedical informatics (as it sometimes had been), and that it was also highly pertinent to the world of health planning and policy. But a major point was my emphasis on the importance of training a cadre of individuals who are deeply skilled in computer science and AI, but also knowledgeable about, and comfortable in, the life science research laboratory and the world of clinical medicine and practice.

3 Evolution of AIM in the Last Decade

A half-century has passed since the earliest explorations of AI in the life sciences and medicine. It is accordingly not surprising that a great deal has changed since the earliest notions were formed regarding the field and its potential. For example, there was an early debate as to whether AI should attempt to understand and simulate human reasoning processes, using methods such as formal logic that were provably correct. Those who extolled this view were dubbed *neats*, whereas the *scruffies* were those who were more concerned with human-like performance, regardless of whether the underlying methods were elegant or were homogenous in the way that they represented and demonstrated the notion of intelligence[8]. Although this distinction continues to exist among AI researchers, the scruffy approach now dominates, and much of the intelligent performance we observe in computers today has little to do with the way that human beings would tackle the same problem. AI has taken on a kind of magical quality, with most people impressed by how smart computers appear to have become but having little insight into how they achieve what they do.

Consider, for example, the long history of AI research in natural language processing and machine learning that led to the 2011 television performance of IBM's *Watson* when it won a Jeopardy competition against two human game champions [13]. People had vague notions that Watson knew a lot, and was very fast, but the long line of research that had led to such performance is not well appreciated. As early as 1959, Arthur Samuel (at IBM) began his exploration of machine learning that resulted in a computer that was able to play checkers, to learn how to win by playing thousands of games against itself after it knew the rules, and eventually to beat human competitors[9]. The perceptron predecessor to neural networks, viewed as a method for recognizing patterns by simulating brain neuronal interconnectivity, was first explored in the 1950s and later popularized in a 1969 book by Minsky and Papert [14], long before neural networks with multiple "deep" hidden layers would have been computationally tractable. Research on natural language processing and text mining also has a long history[10].

We can make similar observations about our everyday acceptance of the speech-understanding capabilities of *Siri* or *Alexa*, both of which owe their performance to decades of AI investigation, such as the influential Hearsay effort at Carnegie Mellon University in the 1960s and 1970s [15]. And when I first arrived at Stanford University in 1970, I was excited when I visited Stanford Research Institute (now known as SRI International) to see the performance of an early wheeled intelligent robot, Shakey, as it solved problems regarding mobility and performed physical tasks[11]. Later, on my first visit to the Stanford AI Laboratory, nestled on a hill above the campus, I drove in and was intrigued to see a sign "Caution:

[8] https://en.wikipedia.org/wiki/Neats_and_scruffies (Accessed Feb 12 2019)

[9] https://gurucul.com/blog/machine-learning (Accessed Feb 12 2019)

[10] https://en.wikipedia.org/wiki/History_of_natural_language_processing (Accessed Feb 12 2019)

[11] https://www.sri.com/work/timeline-innovation/timeline.php?timeline=computing-digital#!&innovation=shakey-the-robot (Accessed Feb 12 2019)

Robot Vehicle". Researchers had begun a line of research on driverless vehicles that has continued for decades[12] and has contributed to the current notion that our Uber may soon arrive without a driver in the front seat[13].

I mention these early AI research activities to stress that what we embrace today did not suddenly appear from a commercial company but was the result of decades of research, typically with slow progress and a visionary set of scientists and government funding agencies who advanced the field. People chatting with intelligent agents in their phones have little insight into what was required to create the technology that they now take for granted. And in the popular media, the phrase "AI" has transformed from being the name of a computer-science subfield into a noun used to describe a sentient but digital being – "there is an AI in my phone!".

There is, of course, also a medical AI version of this story. It is caught up in a confusion of terms that have recently come into vogue. Researchers in AIM, and in biomedical informatics in general, have long struggled with large amounts of data that have pushed the boundaries of what can reasonably be tackled with existing technology. Similarly, they have always developed analytical methods for dealing with such data, combining statistical and computational techniques. Thus "big data" (where "big" is a relative term, reflecting evolution in computing capacity and technology) have always characterized our field. We have always embraced the need for "data analytics" and have developed corresponding methods. However, in recent years, notions of *data science* have emerged as though they are new and have not always been with us. Enthusiasm for new deep machine-learning methods and data science has often failed to acknowledge that those topics are part of a continuum that has been explored by computer scientists, biomedical informaticians, and AIM researchers for decades. In some circles, machine learning and artificial intelligence are seen as nearly synonymous, and it is unfortunate that both academic and government institutions are increasingly creating organizational structures suggesting that biomedical data science and biomedical informatics are separate disciplines or activities.

As previously mentioned, AIM is naturally entwined with other areas of computer science and biomedical informatics. Although there is no chapter on AIM in our textbook of biomedical informatics [16], the field arises throughout the volume in almost every chapter, e.g., those dealing with decision science, cognitive science, human-computer interaction, bioinformatics, natural language processing, decision-support systems, imaging, and the like. As editor of the *Journal of Biomedical Informatics* (JBI), I also have found that a very high percentage of current submissions deal with AIM topics, and especially with NLP and machine learning applied to health care problems. Many are novel applications of existing techniques[14], but some papers introduce new methods motivated by medical problems, and many of these generalize broadly for diverse applications, including outside of biomedicine and health. Similar phenomena are evident when I peruse other informatics journals, leading to my assertion that AIM-related research is a core element in current informatics investigative activities.

Extensive AIM research activity is now, for the first time, matched by enthusiasm for the field[15] and by AI investment by health care organizations and in industry, including new startup companies. Top medical journals have published editorials or articles extolling the promise of AI [17, 18], and the press is filled with articles on the topic[16]. Expectations are high, and the public seems more accepting of the notion that AI will play a role in the care that they receive.

There are cautionary notes as well. For example, some fear that AI could worsen health disparities [19], and some argue that AI is simply being overhyped and has a long way to go [20]. Some of us fear another AI winter could set in, but on balance the field is proving itself with formal studies as well as anecdotal stories.

4 Taking a Deeper Look

Today the clinical literature is replete with formal studies demonstrating the quality and value of AI methods, especially in the interpretation of large datasets, typically using deep learning approaches. Perhaps the most visible successes are those dealing with image interpretation – radiographs [21], ultrasound scans [22], photographs [23], skin lesions [24], microbiology [25], and the like. The results of some studies have been so impressive that there is even debate in the radiological community about whether radiologists will be replaced by AI systems[17]. I find this unlikely, but the nature of radiologic work and expertise will evolve.

The use of deep-learning techniques for image interpretation, and their gradual insertion into routine use, are similar to what the medical world has already accepted in the area of electrocardiographic (ECG) interpretation. In much of the world, ECG reports are now generated by computers and then checked and approved by a cardiologist. The ordering physicians can still review the ECG themselves and decide whether they agree with the interpretation, but there is broad recognition that it is more efficient, and generally just as accurate (or more so, given issues of human fatigue and distractibility), to have a computer perform the initial

[12] http://forum.stanford.edu/wiki/index.php/Timeline_Archive (Accessed Feb 12 2019).

[13] https://www.bbc.com/news/business-45324753 (Accessed Feb 12 2019).

[14] Such papers do not match JBI's editorial policies and we refer them elsewhere for possible publication.

[15] https://www.healthitoutcomes.com/doc/the-year-of-machine-learning-and-ai-in-healthcare-0001 (Accessed Feb 13 2019)

[16] See, for example, the footnotes to the first paragraph in this paper.

[17] "If you think AI will never replace radiologists—you may want to think again." https://www.radiologybusiness.com/topics/artificial-intelligence/if-you-think-ai-will-never-replace-radiologists-you-may-want-think. (Accessed Feb 13 2019)

"Why AI will not replace radiologists." https://towardsdatascience.com/why-ai-will-not-replace-radiologists-c7736f2c7d80. (Accessed Feb 13 2019).

analysis and to generate the report. The arguments would seem to be even stronger in the case of radiologic interpretation, where there is growing evidence that, on average, the AI interpretation will be more accurate. Of course, as in the ECG example, physician review and approval are likely to be required for the foreseeable future.

One issue that arises frequently in the setting of deep learning systems is their lack of transparency: the scoring algorithms cannot easily be translated into explanations regarding the basis for a program's decision. Government research agencies realize that this is an ongoing issue[18], and, given the nature of medicine and the independence of clinicians, explanation has long been viewed as an important element in an AIM system. For programs that use machine learning methods to perform signal or image analysis, however, such explanations may be less pertinent. The ECG example shows that clinicians will accept such systems if they can independently validate the decision offered by a program.

A colleague and I have recently argued that the situation may be different when an AIM system is offering diagnostic or therapeutic advice through a direct interaction with the clinician [26]. We have known for decades that it is especially challenging to motivate clinicians to use a decision-support tool when they must interact with the program themselves (and not simply receive a computer-generated report). Although some colleagues have argued that explanation is not a crucial requirement, now that the use of computers is more clinically accepted, my own experience would suggest otherwise. I have found that clinicians use explanation capabilities less frequently as they became comfortable with an advisory tool, but they want to know that the capability exists in case they have questions. This observation led me to argue in the recent article [26] that "black boxes are unacceptable" when a decision-support tool is designed for direct use by a clinician.

The need for explanation was part of the reason why earlier AIM systems emphasized the "knowledge is power" concept [27], which for decades became a kind of mantra for the expert systems research community. I believe that the true power of our deep-learning applications will occur when we have better approaches for merging the analytical power of machine learning with an explicit encoding of the relevant domain semantics. It is one of the reasons that my recent article [26] stressed that "relevance and insight are essential", explaining that a decision support system "should reflect an understanding of the pertinent domain and the kinds of questions with which clinicians are likely to want assistance."

I also believe that we must counter all suggestions that AIM systems are being designed to make the therapeutic or management decisions. We must stress that it is the synergy between user and machine, and not the machine alone, that properly leverages the unique skills of both participants. It is through this kind of partnership that the care of patients will improve and flourish as AIM systems became routinely integrated into care settings.

5 Moving Forward

With new companies forming, new products being released, and established firms adding major AIM elements to their business strategy, one could ask how the AIM research community might best interface with these new kinds of activities. In the days when there was almost no commercial interest in medical AI, the work was largely based in academia, supported as research, and implemented and evaluated in affiliated clinical settings. Today there is an enhanced potential for partnerships between those in the AIM research community and those involved with the commercialization of the technology.

On the one hand, there should be opportunities for joint research projects, largely academically based, in which industry and academic AIM groups combine their unique skills and resources to develop new methods and systems. There is an existing model for this kind of collaboration in the computer-science community, where firms financially sponsor joint research, students, and even professorships. It is time to forge more such relationships between members of the biomedical informatics academic community and their counterparts in industry.

But there is another way in which the AIM community, and others in academic biomedical informatics, can work proactively and collaboratively with industry. We can take an example from the pharmaceutical industry, which long ago realized that the world of clinical medicine demands evidence that is rigorously developed and published in peer-reviewed journals. If such research were carried out by the company itself, there would be inevitable concerns about conflicts of interest that might taint the study design, its implementation, or the subsequent analysis. As a result, the pharmaceutical firms often collaborate with (and provide funding for) academic medical groups and groups of practitioners who take the lead in the implementation, data gathering, analysis, and publication of the results.

By analogy, the budding AIM industry needs to take seriously the importance of rigorous evaluation of their products and thus the importance of collaborations with independent groups, typically in academia, to carry out the relevant trials. In the case of clinical systems, this will mean industry helping to fund studies that establish a product's safety, validity, reproducibility, usability, and reliability [26]. There is accordingly great synergy that matches the needs and interests of the AIM industry as it evolves and those of the academic AIM and biomedical informatics communities.

I close, as I often do, with an appeal for more training of talented individuals to work in both the academic and industry sides of this AIM research relationship. Due to a paucity of AIM graduates, too many companies have had to settle for hiring individuals with computer-science talent but limited understanding of the culture of medicine or the details of practice settings. The insertion of even one well-trained biomedical informatics scientist into such settings can radically influence the way a product develops, evaluations are performed, or corporate strategies are defined. We could debate the meaning of "well-trained", but at a minimum I believe it implies solid technical skills in computer science and AI, coupled with a deep understanding of medicine and clinical care.

[18] https://www.darpa.mil/program/explainable-artificial-intelligence (Accessed Feb 13 2019)

The future of AIM is bright, building on the remarkable transformation in technology, computing, medicine, and biology over the past half-century. Terminology may change, and experiments will continue to falter along the way, but consideration of the steady progress of five decades is inspiring, even if the rate of progress has been slower than we might have hoped or predicted. We should learn from that experience and avoid predictions today that are ill-considered or that fail to appreciate the complexity of the task that we have undertaken and that will continue to challenge us – often in ways that will surprise us.

References

1. Topol EJ. Hi-performance medicine: the convergence of human and artificial intelligence. Nat Med 2019;25:44-56.
2. Shortliffe EH. The adolescence of AI in Medicine: Will the field come of age in the 90s? Artif Intell Med 1993;5:93-106.
3. Patel VL, Shortliffe EH, Stefanelli M, Szolovits P, Berthold MR, Bellazzi R, et al. The coming of age of artificial intelligence in medicine. Artif Intell Med 2009;46:5-17.
4. Buchanan BG, Sutherland G, Feigenbaum E. Heuristic DENDRAL: A program for generating explanatory hypotheses in organic chemistry. In: Meltzer B, Michie D, editors. Machine Intelligence 4. Edinburgh, UK: Edinburgh Univ Press; 1969. P. 209-54.
5. Sridharan NS. (guest editor). Special issue on artificial intelligence in medicine. Artif Intell 1978;11:1.
6. Clancey WJ, Shortliffe EH. Readings in Medical Artificial Intelligence. Reading, MA: Addison-Wesley; 1984. Available online at http://www.shortliffe.net/Clancey-Shortliffe-1984/Readings%20Book.htm. [Accessed Feb 10 2019].
7. Freiherr G. The seeds of artificial intelligence: SUMEX-AIM. US GPO; DHEW publication no. (NIIH) 80-2071. Washington, DC: US Dept of Health Education, and Welfare, Public Health Service, National Institutes of Health; 1980. Available online at https://exhibits.stanford.edu/feigenbaum/catalog/kd321pr5751. [Accessed Feb 10 2019].
8. Shortliffe EH. Computer-Based Medical Consultations: MYCIN. New York: American Elsevier; 1976. Available online at http://www.shortliffe.net/Shortliffe-1976/MYCIN%20thesis%20Book.htm. (Accessed Feb 10 2019).
9. Buchanan BG, Shortliffe EH. Rule-Based Expert Systems: The MYCIN Experiments of the Stanford Heuristic Programming Project. Reading, MA: Addison-Wesley; 1984. Available online at https://aitopics.org/misc/rule-based-expert-systems-mycin-experiments-stanford-heuristic-programming-project. [Accessed Feb 10 2019].
10. Herzenberg L, Rindfleisch TC, Herzenberg L. Joshua Lederberg: The Stanford Years (1958-1978). Annu Rev Genet 2008;42:19-25.
11. Ohno-Machado, L. Research on machine learning issues in biomedical informatics modeling (editorial for special issue). J Biomed Inform 2004;37:221-3.
12. Ohno-Machado L. Clinical machine learning (editorial for special issue). J Biomed Inform 2005;38:345-6.
13. Markoff J. Computer wins on 'Jeopardy': Trivial, it's not. The New York Times, February 11, 2011. Available online at https://www.nytimes.com/2011/02/17/science/17jeopardy-watson.html. [Accessed Feb 12 2019].
14. Minsky M, Papert S. Perceptrons: An Introduction to Computational Geometry. Cambridge, MA: MIT Press; 1987 (originally published in 1969).
15. Reddy DR, Erman LD, Fennell RO, Neely RB. The hearsay speech understanding system: An example of the recognition process. Proceedings of the 3rd International Joint Conference on Artificial Intelligence. San Francisco, CA: Morgan Kaufmann Publishers; 1973. p. 185-93.
16. Shortliffe EH, Cimino JJ, editors. Biomedical Informatics: Computer Applications in Health Care and Biomedicine. London: Springer; 2014.
17. Hinton G. Deep learning—a technology with the potential to transform health care. JAMA 2018;320(11):1101-2.
18. Liang H, Tsui BY, Ni H, Valentim CCS, Baxter SL, Liu G, Cai W, et al. Evaluation and accurate diagnoses of pediatric diseases using artificial intelligence. Nat Med 2019 Feb 11. doi: 10.1038/s41591-018-0335-9.
19. Khullar D. A.I. Could Worsen Health Disparities. The New York Times, January 31, 2019. Available online at https://www.nytimes.com/2019/01/31/opinion/ai-bias-healthcare.html. [Accessed Feb 13 2019].
20. Marcus G. Artificial Intelligence is Stuck. Here's How to Move it Forward. The New York Times, July 29, 2017. Available online at https://www.nytimes.com/2017/07/29/opinion/sunday/artificial-intelligence-is-stuck-heres-how-to-move-it-forward.html. [Accessed Feb 13 2019].
21. Mayo RC, Leung J. Artificial intelligence and deep learning – Radiology's next frontier? Clin Imaging 2018;49:87-88.
22. Ko SY, Lee JH, Yoon JH, Na H, Hong E, Han K, et al. A deep convolutional neural network for the diagnosis of thyroid nodules on ultrasound. Head Neck 2019 Feb 4. doi: 10.1002/hed.25415.
23. Ting DSW, Cheung CY, Lim G, Tan GSW, Quang ND, Gan A, et al. Development and validation of a deep learning system for diabetic retinopathy and related eye diseases using retinal images from multiethnic populations with diabetes. JAMA 2017 Dec 12;318(22):2211-23.
24. Han SS, Kim MS, Lim W, Park GH, Park I, Chang SE. Classification of the clinical images for benign and malignant cutaneous tumors using a deep learning algorithm. J Invest Dermatol 2018 Jul;138(7):1529-38.
25. Smith KP, Kang AD, Kirby JE. Automated interpretation of blood clulture gram stains by use of a deep convolutional neural network. J Clin Microbiol 2018;56(3).
26. Shortliffe EH, Sepúlveda MJ. Clinical decision support in the era of artificial intelligence. JAMA 2018;320(21):2199-200.
27. Hendler J, Feigenbaum EA. Knowledge is Power: The Semantic Web Vision. In: Zhong N, Yao Y, Liu J, Ohsuga S, editors. Web Intelligence: Research and Development: WI 2001. Lecture Notes in Computer Science, vol 2198. Berlin, Heidelberg: Springer; 2001.

Correspondence to:
E. H. Shortliffe
272 W 107th St Apt 5B
New York, New York 10025-7833
USA
Cell: +1-917-640-0933
E-mail: ted@shortliffe.net; ehs79@columbia.edu