JAMA Health Forum™

What Artificial Intelligence Means for Health Care

David M. Cutler, PhD

Author affiliations and article information are listed at the end of this article.

The artificial intelligence (AI) revolution has started in earnest. Barely a day goes by without learning of new ways to use AI, including in health care. Studies show that AI chatbots can pass medical licensing board examinations, provide second opinions, and display more compassion than physicians.[1-3] What does this mean for physicians?

No one knows how AI will ultimately affect medicine, but theory and experience in other industries provide some guidance. Here are 5 observations about the possible effects of AI on medicine.

First, AI is likely to substitute for rote activities that humans currently perform, such as routine office work: billing, appointment scheduling, and facility management. Currently, these tasks are people intensive; with AI, the need for and cost of office staff can be reduced.[4] Recently, colleagues and I estimated that savings from AI in this domain could range from $200 billion to $360 billion annually, about 35% of which would be administrative savings.[4]

This substitution will lead to a reduction in health care employment, but it is likely to be gradual as use of AI expands across areas (billing, management, scheduling, and so on). Think not about robots replacing factory workers, but about the gradual decline in office administrative staff over time.

Second, in clinical care, AI is more likely to complement clinicians than substitute for them.[5] Although clinical care generally is not administered in the same routine way as administrative tasks, there are some duties in which AI can be useful, such as scanning laboratory results for abnormalities. But clinical care also involves more subtle aspects that AI cannot yet mimic. Is this a patient that typically reports a good deal of pain? Does the patient look more confused than usual? Clinicians will still need to combine structured and unstructured—and even unrecorded (eg, physical appearance, sound of the voice, hesitancy in speaking)—data.

Learning how to integrate human decision-making with software is extremely important. It would be quite frustrating for physicians to arrive at a diagnosis and treatment plan only for AI to upend them with a critique and alternative options. At the same time, receiving reams of raw output from AI programs will overwhelm physicians, who already suffer from "click fatigue."[6] Designing effective systems for interactions between humans and AI is thus essential.

Third, it is particularly important to develop AI applications that enhance efficiency by enabling less expensive monitoring, diagnosis, and personnel needs. The most expensive medical care occurs in institutions: hospitals and postacute facilities. Patients often receive such care because they need continuous monitoring. By making remote monitoring easier, AI can help move some of this care into the home or a step-down observation unit.

The cost of diagnosing illness may become cheaper as well. For example, eligibility for clinical trials of Alzheimer disease therapies often requires a diagnostic positron emission tomographic scan or analysis of cerebrospinal fluid, both of which are complex and expensive.[7] Using AI to analyze blood biomarkers, perhaps in combination with less expensive brain imaging, could replace more expensive modalities in clinical research and practice and lead to reductions in high-cost imaging and invasive testing.

Artificial intelligence can also help substitute less expensive clinicians for more expensive ones. Anesthesia is a case in point. Anesthesia administration used to be sufficiently complex that a high level of training was required for all applications. However, standardizing administration protocols

has meant that nurse anesthetists can substitute for anesthesiologists in many routine cases. Artificial intelligence may be the key to such substitutions in other areas of medicine (eg, by guiding mid-level clinicians through appropriate processes).

Fourth, AI algorithms should not just replicate human thinking processes but should aim to exceed them. Humans may make poor or biased decisions, some of which are random, whereas other decisions systematically affect individuals with lower incomes, lower education levels, and racial and ethnic minority groups.[8] Building software code and algorithms that replicate human mistakes and bias is not good enough.[9] Software that predicts the ground truth, not a fallible individual's interpretation of it, is needed.

One step required for this is to increase the quantity and quality of data available. One of the classic uses of machine learning is pattern recognition: is the object in the picture a cat or a dog? Outside health care, pattern recognition algorithms are based on millions of images. In health care, a large quantity of images and data exist, but they are buried in practice-specific electronic medical records. Thus, clinicians have access to limited data and lack computer science expertise, whereas computer scientists have technical skills but not enough access to data. A major priority is to gather large samples of data that are based on ground truth rather than just perceptions of truth.

Fifth, it is important to be clear about what AI is not good at. Machine learning can be terrific at finding patterns in data. It can scan clinical trial outputs and identify subgroups with larger than average treatment effects. However, this is more akin to hypothesis generation than proof of differences because there will always be some subgroups with a larger than average response in any data set. Making causal determinations will still require classic hypothesis testing—either from clinical trials or using quasi-experimental settings in the real world.

Historically, new technology has often been extremely beneficial for patients and clinicians. The entire clinical enterprise is built on the ability of clinicians to harness data and devices to improve patients' health. Hopefully, this will be true about AI as well, with economic benefits to boot. But making this happen will require conscious effort and planning; it cannot be left to chance.

Corresponding Author: David M. Cutler, PhD, Department of Economics, Harvard University, 1805 Cambridge St, Cambridge, MA 02138 (dcutler@fas.harvard.edu).

Author Affiliation: Department of Economics and Kennedy School of Government, Harvard University, Cambridge, Massachusetts.

REFERENCES

1. Kung TH, Cheatham M, Medenilla A, et al. Performance of ChatGPT on USMLE: potential for AI-assisted medical education using large language models. *PLOS Digit Health*. 2023;2(2):e0000198. doi:10.1371/journal.pdig.0000198

2. Lee P, Bubeck S, Petro J. Benefits, limits, and risks of GPT-4 as an AI chatbot for medicine. *N Engl J Med*. 2023;388(13):1233-1239. doi:10.1056/NEJMsr2214184

3. Ayers JW, Poliak A, Dredze M, et al. Comparing physician and artificial intelligence chatbot responses to patient questions posted to a public social media forum. *JAMA Intern Med*. 2023;183(6):589-596. doi:10.1001/jamainternmed.2023.1838

4. Sahni NR, Stein G, Zemmel R, Cutler DM. The potential impact of artificial intelligence on healthcare spending. Published January 2023. Accessed June 3, 2023. https://www.nber.org/papers/w30857

5. Mello MM, Guha N. ChatGPT and physicians' malpractice risk. *JAMA Health Forum*. 2023;4(5):e231938. doi:10.1001/jamahealthforum.2023.1938

6. Jamoom EW, Heisey-Grove D, Yang N, Scanlon P. Physician opinions about EHR use by EHR experience and by whether the practice had optimized its EHR use. *J Health Med Inform*. 2016;7(4):1000240.

7. Arias JJ, Phillips KA, Karlawish J. Developing an economic and policy research agenda for blood biomarkers of neurodegenerative diseases. *JAMA Health Forum*. 2021;2(7):e211428. doi:10.1001/jamahealthforum.2021.1428

8. Pierson E, Cutler DM, Leskovec J, Mullainathan S, Obermeyer Z. An algorithmic approach to reducing unexplained pain disparities in underserved populations. *Nat Med*. 2021;27(1):136-140. doi:10.1038/s41591-020-01192-7

9. Obermeyer Z, Powers B, Vogeli C, Mullainathan S. Dissecting racial bias in an algorithm used to manage the health of populations. *Science*. 2019;366(6464):447-453. doi:10.1126/science.aax2342