

University of Michigan Law School

University of Michigan Law School Scholarship Repository

Other Publications

Faculty Scholarship

2019

Risks and Remedies for Artificial Intelligence in Healthcare

W. Nicholson Price II

University of Michigan School of Law, wnp@umich.edu

Available at: <https://repository.law.umich.edu/other/240>

Follow this and additional works at: <https://repository.law.umich.edu/other>



Part of the [Artificial Intelligence and Robotics Commons](#), [Health Law and Policy Commons](#), and the [Science and Technology Law Commons](#)

Recommended Citation

Price, W. Nicholson, II. *Risks and Remedies for Artificial Intelligence in Healthcare*. Brookings Institution, 2019. (Report from The Brookings Institution's Artificial Intelligence and Emerging Technology Initiative.)

This Report is brought to you for free and open access by the Faculty Scholarship at University of Michigan Law School Scholarship Repository. It has been accepted for inclusion in Other Publications by an authorized administrator of University of Michigan Law School Scholarship Repository. For more information, please contact mlaw.repository@umich.edu.

BROOKINGS

Report

Risks and remedies for artificial intelligence in health care

W. Nicholson Price II Thursday, November 14, 2019

Editor's Note:

This report from The Brookings Institution's Artificial Intelligence and Emerging Technology (AIET) Initiative is part of "[AI Governance](#)," a series that identifies key governance and norm issues related to AI and proposes policy remedies to address the complex challenges associated with emerging technologies.

Introduction

Artificial intelligence (AI) is rapidly entering health care and serving major roles, from automating drudgery and routine tasks in medical practice to managing patients and medical resources. As developers create AI systems to take on these tasks, several risks and challenges emerge, including the risk of injuries to patients from AI system errors, the risk to patient privacy of data acquisition and AI inference, and more. Potential solutions are complex but involve investment in infrastructure for high-quality, representative data; collaborative oversight by both the Food and Drug Administration and other health-care actors; and changes to medical education that will prepare providers for shifting roles in an evolving system.

Potential benefits

Although the field is quite young, AI has the potential to play at least four major roles in the health-care system:^[1]

Pushing boundaries of human performance. The flashiest use of medical AI is to do things that human providers—even excellent ones—cannot yet do. For instance, Google Health has developed a program that can predict the onset of acute kidney injury up to

two days before the injury occurs; compare that to current medical practice, where the injury often isn't noticed until *after* it happens.^[2] Such algorithms can improve care beyond the current boundaries of human performance.

“The flashiest use of medical AI is to do things that human providers—even excellent ones—cannot yet do.”

Democratizing medical knowledge and excellence. AI can also share the expertise and performance of specialists to supplement providers who might otherwise lack that expertise. Ophthalmology and radiology are popular targets, especially because AI image-analysis techniques have long been a focus of development. Several programs use images of the human eye to give diagnoses that otherwise would require an ophthalmologist. Using these programs, general practitioner, technician, or even a patient can reach that conclusion.^[3] Such democratization matters because specialists, especially highly skilled experts, are relatively rare compared to need in many areas.

Automating drudgery in medical practice. AI can automate some of the computer tasks that take up much of medical practice today. Providers spend a tremendous amount of time dealing with electronic medical records, reading screens, and typing on keyboards, even in the exam room.^[4] If AI systems can queue up the most relevant information in patient records and then distill recordings of appointments and conversations down into structured data, they could save substantial time for providers and might increase the amount of facetime between providers and patients and the quality of the medical encounter for both.

Managing patients and medical resources. Finally, and least visibly to the public, AI can be used to allocate resources and shape business. For instance, AI systems might predict which departments are likely to need additional short-term staffing, suggest which

of two patients might benefit most from scarce medical resources, or, more controversially, identify revenue-maximizing practices.

Risks and challenges

While AI offers a number of possible benefits, there also are several risks:

Injuries and error. The most obvious risk is that AI systems will sometimes be wrong, and that patient injury or other health-care problems may result. If an AI system recommends the wrong drug for a patient, fails to notice a tumor on a radiological scan, or allocates a hospital bed to one patient over another because it predicted wrongly which patient would benefit more, the patient could be injured. Of course, many injuries occur due to medical error in the health-care system today, even without the involvement of AI. AI errors are potentially different for at least two reasons. First, patients and providers may react differently to injuries resulting from software than from human error. Second, if AI systems become widespread, an underlying problem in one AI system might result in injuries to thousands of patients—rather than the limited number of patients injured by any single provider’s error.

Data availability. Training AI systems requires large amounts of data from sources such as electronic health records, pharmacy records, insurance claims records, or consumer-generated information like fitness trackers or purchasing history. But health data are often problematic. Data are typically fragmented across many different systems. Even aside from the variety just mentioned, patients typically see different providers and switch insurance companies, leading to data split in multiple systems and multiple formats. This fragmentation increases the risk of error, decreases the comprehensiveness of datasets, and increases the expense of gathering data—which also limits the types of entities that can develop effective health-care AI.

Privacy concerns. Another set of risks arise around privacy.^[5] The requirement of large datasets creates incentives for developers to collect such data from many patients. Some patients may be concerned that this collection may violate their privacy, and lawsuits have been filed based on data-sharing between large health systems and AI developers.^[6] AI

could implicate privacy in another way: AI can predict private information about patients even though the algorithm never received that information. (Indeed, this is often the goal of health-care AI.) For instance, an AI system might be able to identify that a person has Parkinson’s disease based on the trembling of a computer mouse, even if the person had never revealed that information to anyone else (or did not know). Patients might consider this a violation of their privacy, especially if the AI system’s inference were available to third parties, such as banks or life insurance companies.

Bias and inequality. There are risks involving bias and inequality in health-care AI. AI systems learn from the data on which they are trained, and they can incorporate biases from those data. For instance, if the data available for AI are principally gathered in academic medical centers, the resulting AI systems will know less about—and therefore will treat less effectively—patients from populations that do not typically frequent academic medical centers. Similarly, if speech-recognition AI systems are used to transcribe encounter notes, such AI may perform worse when the provider is of a race or gender underrepresented in training data.^[7]

“Even if AI systems learn from accurate, representative data, there can still be problems if that information reflects underlying biases and inequalities in the health system.”

Even if AI systems learn from accurate, representative data, there can still be problems if that information reflects underlying biases and inequalities in the health system. For example, African-American patients receive, on average, less treatment for pain than white patients;^[8] an AI system learning from health-system records might learn to suggest lower doses of painkillers to African-American patients even though that decision reflects systemic bias, not biological reality. Resource-allocation AI systems could also exacerbate inequality by assigning fewer resources to patients considered less desirable or less profitable by health systems for a variety of problematic reasons.

Professional realignment. Longer-term risks involve shifts in the medical profession. Some medical specialties, such as radiology, are likely to shift substantially as much of their work becomes automatable. Some scholars are concerned that the widespread use of AI will result in decreased human knowledge and capacity over time, such that providers lose the ability to catch and correct AI errors and further to develop medical knowledge.^[9]

The nirvana fallacy. One final risk bears mention. AI has the potential for tremendous good in health care. The nirvana fallacy posits that problems arise when policymakers and others compare a new option to perfection, rather than the status quo. Health-care AI faces risks and challenges. But the current system is also rife with problems. Doing nothing because AI is imperfect creates the risk of perpetuating a problematic status quo.

Possible solutions

There are several ways we can deal with possible risks of health-care AI:

Data generation and availability. Several risks arise from the difficulty of assembling high-quality data in a manner consistent with protecting patient privacy. One set of potential solutions turns on government provision of infrastructural resources for data, ranging from setting standards for electronic health records to directly providing technical support for high-quality data-gathering efforts in health systems that otherwise lack those resources. A parallel option is direct investment in the creation of high-quality datasets. Reflecting this direction, both the United States' All of Us initiative and the U.K.'s BioBank aim to collect comprehensive health-care data on huge numbers of individuals. Ensuring effective privacy safeguards for these large-scale datasets will likely be essential to ensuring patient trust and participation.

Quality oversight. Oversight of AI-system quality will help address the risk of patient injury. The Food and Drug Administration (FDA) oversees some health-care AI products that are commercially marketed. The agency has already cleared several products for market entry, and it is thinking creatively about how best to oversee AI systems in health. However, many AI systems in health care will not fall under FDA's purview, either because they do not perform medical functions (in the case of back-end business or resource-allocation AI) or because they are developed and deployed in-house at health systems

themselves—a category of products FDA typically does not oversee. These health-care AI systems fall into something of an oversight gap. Increased oversight efforts by health systems and hospitals, professional organizations like the American College of Radiology and the American Medical Association, or insurers may be necessary to ensure quality of systems that fall outside the FDA’s exercise of regulatory authority.^[10]

“A hopeful vision is that providers will be enabled to provide more-personalized and better care. ... A less hopeful vision would see providers struggling to weather a monsoon of uninterpretable predictions and recommendations from competing algorithms.”

Provider engagement and education. The integration of AI into the health system will undoubtedly change the role of health-care providers. A hopeful vision is that providers will be enabled to provide more-personalized and better care, freed to spend more time interacting with patients as humans.^[11] A less hopeful vision would see providers struggling to weather a monsoon of uninterpretable predictions and recommendations from competing algorithms. In either case—or in any option in-between—medical education will need to prepare providers to evaluate and interpret the AI systems they will encounter in the evolving health-care environment.

The Brookings Institution is a nonprofit organization devoted to independent research and policy solutions. Its mission is to conduct high-quality, independent research and, based on that research, to provide innovative, practical recommendations for policymakers and the public. The conclusions and recommendations of any Brookings publication are solely those of its author(s), and do not reflect the views of the Institution, its management, or its other scholars.

Microsoft provides support to The Brookings Institution's Artificial Intelligence and Emerging Technology (AIET) Initiative, and Google provides general, unrestricted support to the Institution. The findings, interpretations, and conclusions in this report are not influenced by any donation. Brookings recognizes that the value it provides is in its absolute commitment to quality, independence, and impact. Activities supported by its donors reflect this commitment.

Report Produced by **Center for Technology Innovation**

Footnotes

1. 1 W. Nicholson Price II, Artificial intelligence in the medical system: four roles for potential transformation, 18 Yale J. Health Pol'y L. & Ethics (forthcoming 2019), 21 Yale J.L. & Tech. (forthcoming 2019), https://papers.ssrn.com/abstract_id=3341692.
2. 2 Nenad Tomašev et al., A clinically applicable approach to continuous prediction of future acute kidney injury, *Nature* 572: 116-119 (2019).
3. 3 See yediagnosis.net for examples.
4. 4 Lauren Block et al., In the wake of the 2003 and 2011 duty hours regulations, how do internal medicine interns spend their time?, *J. Gen. Intern. Med.* 28(8):1042-1047 (2013).
5. 5 W. Nicholson Price II & I. Glenn Cohen, Privacy in the age of medical big data, *Nature Medicine* 25:37-43 (2019).
6. 6 I. Glenn Cohen & Michelle M. Mello, Big data, big tech, and protecting patient privacy, *JAMA* (published online Aug. 9, 2019), <https://jamanetwork.com/journals/jama/fullarticle/2748399>.
7. 7 Joan Palmiter Bajorek, Voice recognition still has significant race and gender biases, *Harvard Bus. Rev.* (May 10, 2019), <https://hbr.org/2019/05/voice-recognition-still-has-significant-race-and-gender-biases>.
8. 8 Monika K. Goyal et al., Racial disparities in pain management of children with appendicitis in emergency departments, *JAMA Pediatrics* 169(11):996-1002 (2015).
9. 9 A. Michael Froomkin et al., When AIs Outperform Doctors: The Dangers of a Tort-Induced Over-Reliance on Machine Learning, 61 *Ariz. L. Rev.* 61:33 (2019)
10. 10 W. Nicholson Price II, Regulating black-box medicine, *Mich. L. Rev.* 116(3):421-474 (2017).
11. 11 Eric Topol, *Deep Medicine*, (Basic Press 2019).