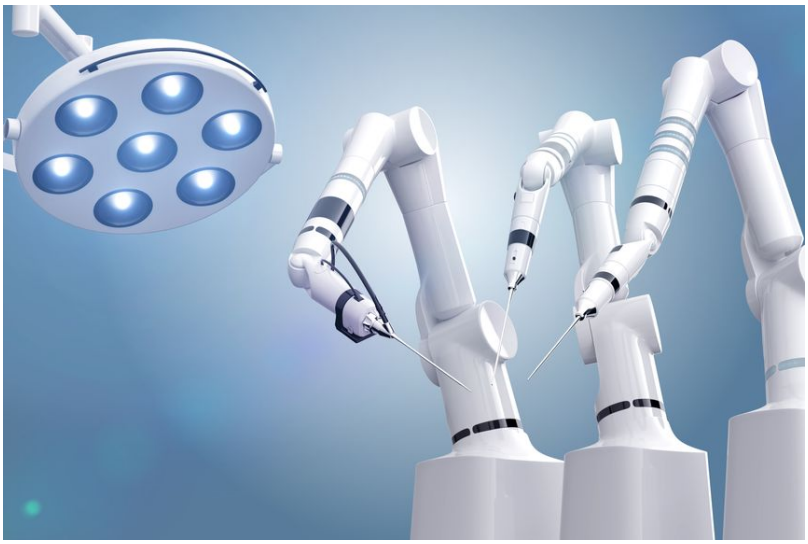
JOURNAL REPORTS: TECHNOLOGY

The Ethical Dilemmas AI Poses for Health Care

Artificial intelligence promises all sorts of advances for medicine. And all sorts of concerns.



There is a danger that machine-learning algorithms can exacerbate human biases and discrimination. PHOTO: GETTY IMAGES/ISTOCKPHOTO

By Lisa Ward

Updated Oct. 14, 2019 7:02 pm ET

Could a machine-learning algorithm diagnose your next illness?

Artificial intelligence can make diagnoses from digitized images such as mammograms and diabetic retinal scans. More sophisticated interventions might also be possible someday: algorithms that guide robots through surgery, for example, or even help restore motor control in paralyzed patients.

Algorithms learn by combing through massive amounts of electronic health records, insurance claims, medical research and other sources. They use the data to identify new patterns, which then become building blocks for new capabilities, such as diagnoses. Some experts believe AI can remake health care, for instance, by reducing the rate of medical errors and creating a repository of medical knowledge that can be consulted widely and understood easily by clinicians.

JOURNAL REPORT

- Read more at WSJ.com/AIreport

MORE IN ARTIFICIAL INTELLIGENCE

- What Will an AI World Be Like?
- Creating a Robot-Human Bond
- AI Learns to Write Fake News
- Relationship Help From an Algorithm

There are big concerns, too. How the algorithms arrive at particular conclusions, for example, is often unclear to people. There is also a danger that machine-learning algorithms—which typically operate within the limitations of whatever data they are fed—can exacerbate human biases and discrimination. Some algorithms trained on data from Caucasian patient populations have done a poor job of making predictions for nonwhite patients.

The Wall Street Journal hosted a conversation by email about these issues with I. Glenn Cohen, a professor at Harvard Law School and director of its Petrie-Flom Center for Health Law Policy, Biotechnology and Bioethics; Ziad Obermeyer, acting associate professor of health policy and management at the University of California, Berkeley; and Danton Char, assistant professor of anesthesiology, perioperative and pain medicine at Stanford University Medical Center. An edited transcript follows.

WSJ: *What are your biggest hopes for AI and machine learning in health care?*

PROF. COHEN: My biggest hope is that AI and machine learning usher in safer, more effective and more cost-effective ways of treating patients. We are already seeing some attempts to achieve this, mostly in imaging—the low-hanging fruit for machine learning.

DR. CHAR: My hope is that it will allow a breadth of perspective not yet captured by either the anecdotal experience of the individual clinician or individual research studies, and that it will increase analytical speed and power to assess interventions or treatment choices.

DR. OBERMEYER: When I work in the emergency department, learning about a new patient feels like drowning in information. What electronic medical records have done is convert a large, inaccessible stack of paper charts into a never-ending stack of virtual notes and results. It's not a good task for humans; we just don't have the bandwidth to process it, and we waste a lot of time, both consuming and producing this stack.

I. Glenn Cohen

One way I sometimes think about algorithms is as a very efficient compression method: For a given decision—e.g., should I test this patient for heart attack—it extracts all the relevant information from the virtual stack of data and reduces it to one number.

WSJ: *What are some of your fears for AI in medicine?*

DR. CHAR: The first fear is that AI and machine learning may worsen the economic and racial disparities already inherent in U.S. health care. Second, U.S. health care is in a constant tension between profit and delivery of health. Those two agendas rarely line up. Consequently, the values of AI designers or the purchasing administrators are not necessarily the values of the bedside clinician or patient. Those value collisions and tensions are going to be sites of significant ethical conflict.

PROF. COHEN: Biggest fears: that medical AI gets used in practice before it has been well validated for clinical use and produces bad outcomes for patients; or, that even when it is well

tested on data sets in its implementation in real hospital systems, the realities of particular workflows, work styles, etc., produce bad patient outcomes.

Alternatively, that medical AI is very successful and produces really good outcomes but becomes a luxury item in patient practice. So medical AI makes those patients with excellent care get even better care but does nothing to democratize expertise.

The world of medical AI requires patients to give up a lot of privacy and increases the chance that they may be more easily re-identifiable, with the attendant stigma and material consequences.

Ziad Obermeyer

DR. OBERMEYER: Over the next few years, as Glenn [Cohen] notes, we'll see many of the easy fixes that we've been promised fall flat—algorithms that are found to do little or do harm when they are rigorously evaluated. So, my worry is that these problems will make us want to throw the baby out with the bathwater, to give up on the huge upside of algorithms for medicine. It's more than just reducing cost or automating some tasks—there is a very real chance that we can fundamentally transform not just the delivery of medicine but also the science underlying it. And, if we get it right, that can do a lot of good in the world.

WSJ: *Could AI undermine physicians' discretion?*

PROF. COHEN: One interesting question has to do with how, in the future, AI will interact with health-insurance reimbursement. Physicians may retain the discretion to deviate from AI recommendations. But that may not matter if the insurer is allowed to say we will only reimburse for what the AI recommends.

DR. CHAR: AI systems might take on an authority they don't deserve. There is a likely scenario where AI systems move into a role of controlling clinicians' decisions and workflow, as the electronic medical records prompts and warnings are already starting to do. This is great when it can be used to demonstrably reduce errors and improve health outcomes; not so great if the chief utility is improving profit or finessing evaluation metrics.

Deviating or challenging algorithmic recommendation is going to be challenging. Cathy O'Neil [author of "Weapons of Math Destruction"] has written about this phenomenon in non-health-care contexts: that not only is there a tendency to overtrust an algorithm's output, but, when individuals dispute the output, they often have to furnish considerably more and better-quality data to rebut the algorithm than the data or evidence the algorithm was built on.

PROF. COHEN: As a nonmedical doctor, but a patient, I can tell you the thing I worry about most is that the current form of medical education is not preparing physicians to practice medicine in an AI-augmented environment; that they are not taught enough data science to be critical users rather than thoughtless users of the technology. I also think we need a lot of "AI Bioethics," but happily many more in my field are turning their attention to it.

DR. OBERMEYER: I couldn't agree more. Everyone agrees that algorithms are going to be a big part of medicine in the future, but we aren't developing the human-capital pipeline against that goal. We don't see it reflected in our pre-med requirements, or medical-school curricula.

DR. CHAR: Experts may focus on certain applications or health-care uses but often lack expertise in other areas, like bioinformatics, clinical contexts and health-care management. So, while individuals may be experts in machine learning or oncology, there is no clear expertise yet in machine learning for oncology.

WSJ: *You are familiar with the "black box" problem: When AI programs can't explain exactly how they come up with their results, should we hold back on adopting it?*

DR. CHAR: I don't know that algorithms for health care need to be transparent, but they need to be explainable or at least auditable. For example, clinicians order MRI scans, but they can't necessarily explain how an MRI works or take apart an MRI and explain its architecture. But *someone* can, and at a clinical level its output, scans, correlates with medically significant findings. Algorithms will need to be similarly constructed both so an engineer/computer scientist can evaluate that the algorithm is doing what it purports to do, and this needs to be rendered in a clinically interpretable fashion for clinical pathologies.

But, transparency in and of itself may not be necessary. There are many areas of medicine—my own field, anesthesia, for example—where we don't know how a therapy works (in this case how inhaled anesthetic drugs induce the anesthetized state) but we can demonstrate that it reliably produces the desired therapeutic effect and that the effect can be monitored.

Danton Char

DR. OBERMEYER: Exactly right. Medicine uses a lot of things that we don't understand and can't explain. So the last thing we'd want to do is hold algorithms to that standard. We want them to discover and help us learn about things we don't understand. In fact, that's most of the upside.

PROF. COHEN: I am with Danton [Char], that what matters is that we have a good reason to believe that AI in medicine will improve outcomes, that it works—not its mechanism of action, or how it works. But I wonder if the average patient or regulator sees it that way? That is, will they make explainability a necessary condition for implementation/comfort with the technology? And if patients are uncomfortable when exposed to the fact that the physician doesn't know how it works (or, for more opaque algorithms, maybe few or no physicians know), is that something they should keep from the patient?

DR. OBERMEYER: I find it useful to compare it to regulation of drugs, where explainability is not really part of the FDA's criteria. We just care that it works. The millions taking SSRIs [a type of antidepressant] don't seem to mind that neither they nor their doctor have any idea how it works.

WSJ: *How should machine-learning algorithms be tested?*

DR. OBERMEYER: I don't think there's any way around randomized trials. We have a playbook for evaluating new technologies in medicine, and ultimately, an algorithm needs to affect outcomes we care about to be worthwhile. If a pharmaceutical company said, "Trust us, we've run a study in our data and our product works," we would not take that statement at face value; why would we do that when the product is an algorithm?

Medicine is just too complicated for anyone to be sure that something works without running a trial. It's one of the amazing things about medicine, as opposed to many other social sectors.

WSJ: *Should patients be informed when an algorithm was used to determine the best course of care, especially if a doctor can't explain why the algorithm chose a particular action?*

DR. CHAR: A problem we're already starting to consider is whether to disclose mortality predictions [from machine-learning algorithms] for patients and what limits to place on their use. Machine learning can be used to make mortality predictions on patients, such as to gauge who might benefit from a palliative-care consultation. However, mortality predictions can also be co-opted for ethically thornier uses such as guiding clinical choices to influence therapies clinicians might offer based on predicted survival.

DR. OBERMEYER: I'd say this is one area where we don't necessarily need a trial. People want the information and we have an obligation to give it to them, to help them plan their finances, family situation, and a million other things. Now, there's another issue entirely, which is using mortality predictions to optimize end-of-life care—to reduce "waste." That's where a trial would be very useful.

WSJ: *AI has been known to exacerbate human biases, including discrimination against certain groups. How has that played out in health care?*

DR. OBERMEYER: When algorithms learn from us—when they learn to discriminate, to be biased—in a way, it's not the algorithm's fault. It's our own fault. When an automated hiring assistant, trained to screen résumés based on how humans screen résumés, produces bias, that makes us conscious of the bias that was already deeply ingrained in the hiring process. And by making that transparent, it lets us fix both the process and the algorithm. So certainly, if algorithm development is done without taking structural inequalities and bias into account very carefully, we'll get predictably awful results. But exactly for all these reasons, algorithms can be a powerful tool for diagnosing—and then helping to fix—these same problems.

DR. CHAR: In my work, it manifests as this "self-fulfilling prophecy" problem, which is to say certain genetic diagnoses or clinical conditions (extreme prematurity, traumatic brain injury) become known as nonsurvivable conditions. Consequently, clinicians withdraw care or limit interventions in patients with these conditions and so the data continues to describe these

conditions as nonsurvivable or highly morbid, until someone starts to aggressively challenge the prophecy.

My concern is not really with machine learning, but that in a progressively shift-based, time-pressured, financially squeezed clinical-work environment, I'm pessimistic about who is going to have the bandwidth to notice the problems. I haven't seen a machine-learning system that's going to notice, say, that the bacteria don't grow around the mold [Alexander Fleming's discovery that led to the invention of penicillin].

DR. OBERMEYER: I love the penicillin example—yes, that's human learning, in all its mystery and at its finest. But here's another example: Brugada syndrome [a genetic disorder] was discovered because a doctor noticed a strange ECG waveform in a child who had a cardiac arrest. Those are the kinds of facts that, at scale, machine learning can produce far better than humans. And those are the kinds of projects I'm working on now that I'm in some ways most excited about, because they let us bring the tools of machine learning to bear on really important problems in medical discovery.

It's certainly not well suited for all problems, but I think our current discovery pipeline can benefit an enormous amount from these tools.

PROF. COHEN: My bigger fear is about lost opportunities. Some of the best-use cases for AI in medicine are about exporting our best diagnostics to low-resource settings in the U.S. or abroad —to give everyone really high-quality diagnosis without seeing the sub-specialist at the fancy teaching hospital. My worry, though, is that, because of the way health care is paid for, that may not be the attractive use case for many AI developers. And, thus, we may miss some of the most powerful ways to make the world healthier.

Ms. Ward is a writer in Mendham, N.J. She can be reached at reports@wsj.com.

Correction
It was Alexander Fleming whose observation that bacteria wasn't growing around mold in a petri dish led to the invention of penicillin. An earlier version of this article incorrectly stated that it was Louis Pasteur. October 14, 2019.

SHARE YOUR THOUGHTS

What concerns do you have about artificial intelligence in the healthcare arena?Join the conversation below.

- College Rankings
- College Rankings Highlights
- Energy
- Funds/ETFs
- Health Care
- Leadership
- Retirement
- Small Business
- Technology
- Wealth Management