

Slight alterations, a sharper bend to the curve, rearrangement, upheaval, and disruption—innovation can have any number of consequences that vary in degree and kind. We report on 11 innovations in American health care in the pages ahead. We won't play up any of our favorites: You decide which of them will make a big splash, which will fizzle, which will plod along and maybe make a difference in the long haul. And be sure to download and use the Zappar app to view added videos and enhanced graphics.

AI for all that ails American health care. But how smart is that?

The evidence base is wobbly, but artificial intelligence is coming on strong and the hype for it is even stronger. Screening for diabetic retinopathy is an early application. Lingering questions include whether the use of AI will contribute to health care inequities rather than solve them.

It seems like only several years ago, the health care industry was beginning to ask what problems could AI solve. Now, one might wonder what AI isn't trying to fix.

Rock Health, a health technology investment and research company, reports that from 2011 to 2017, 121 digital health companies were showered with \$2.7 billion in venture funding to apply AI to 19 areas, ranging from drug research and development to clinical-decision support to health benefits administration.

Private investors aren't alone in betting on AI development. Earlier this year CMS launched an AI outcomes challenge, offering \$1.65 million total in prizes for AI models that predict hospital and skilled nursing facility readmissions and adverse events.

But AI is in that zig and zag of hope and hype that innovation in health care often travels. Good evidence is in short supply—it's patchy, cited selectively by proponents, still very much growing.

Eric Topol, MD, is hardly a Luddite. The founder and director of Scripps Research Translational Institute and an early adopter of remote patient monitoring, Topol is a garrulous, optimistic presence on Twitter when it comes to AI for genomic and other kinds of basic research. But in a review of algorithms for diagnosis and prognosis published in *Nature Medicine* earlier this year, he hit a sobering, "yes, but" note. "The field is certainly high on promise," he concluded, "and relatively low on data and proof."

It doesn't take much imagination to envision an errant AI system missing diagnoses or ordering unnecessary tests. And even when AI can perform as well as skilled (if imperfect) human experts, will that be good enough? Health technology investor Bob Kocher, MD, doesn't

think computers will be forgiven for mistakes as readily as doctors are, and at least one example outside health care suggests he might be right: Last year Uber suspended self-driving vehicle testing on public roads for nine months and terminated 100 autonomous vehicle operators after a single fatal accident involving a self-driving car.

Intelligent as it may be, AI can't escape the limits of the data it is musing on. AI has the potential to amplify existing disparities in care, rather than sweep them away. And AI will likely have to elbow its way into a health care sector rife with entrenched interests, established workflows, and competing agendas.

AI's applications

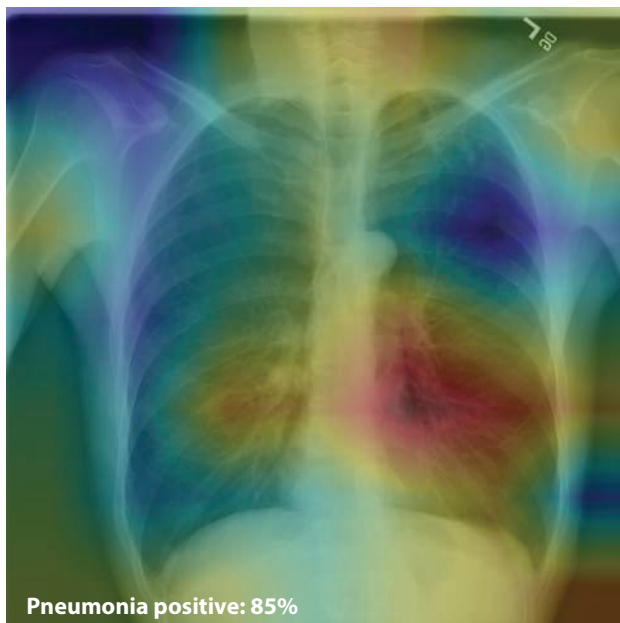
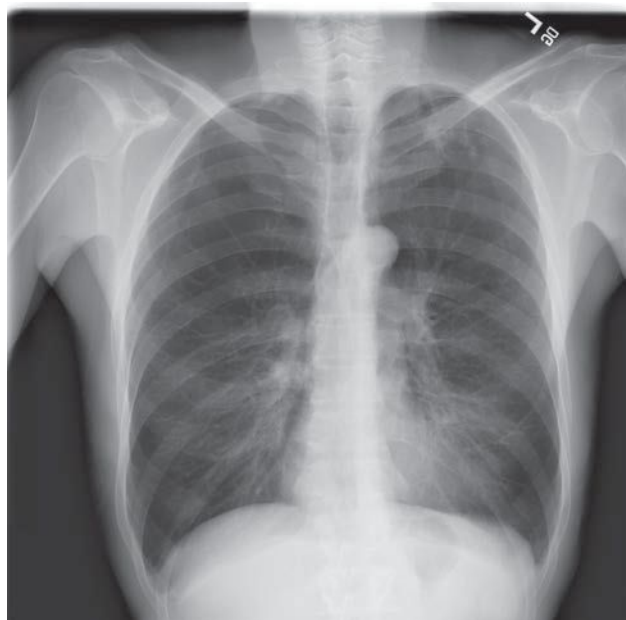
Like almost any trendy topic in health care, AI lacks a codified definition. Vagueness makes it easier to hop on the bandwagon. But a good-enough definition is that AI is technology that can mimic human capabilities like learning and reasoning. Subtypes in the taxonomy include machine learning, natural language processing, and robotics. Machine learning—computer algorithms that learn from data rather than being programmed statically to find patterns and make predictions—is one of the most common AI subtypes in health care. Deep learning—a branch of machine learning that uses a layered structure of algorithms to find patterns in large,

unstructured data sets—is being used to predict clinical outcomes and diagnose diseases. It's essentially a technique that machines can use to learn from data and that is inspired by the way a brain learns to solve problems.

In simplified terms, the current AI applications in American health care fall in two buckets: improving clinical care or streamlining administrative inefficiencies. So far, AI for analyzing images and diagnosing disease has captured the lion's share of attention. But AI is also being used to reduce documentation time for doctors, detect fraud, and automate customer service. Apple Watch's atrial fibrillation detection feature is a beachhead in consumer-oriented AI.

For his balloon-popping review, Topol had searched for prospective studies conducted in real-world clinical settings and published in peer-reviewed journals. He found only eight, including studies evaluating algorithms aiming to detect diabetic retinopathy, wrist fractures in emergency rooms, tiny polyps during colonoscopies, and a couple of other conditions.

Notably, algorithms that predict clinical outcomes—already being used by some payers and providers—didn't make the cut. Blue Cross and Blue Shield of North Carolina and University of Pittsburgh Medical Center have announced they are using or testing AI to predict hospital readmissions. Topol reviewed many studies evaluating algorithms aimed at predicting clini-



In this demonstration from the Stanford ML Group, researchers built a 121-layer convolutional neural network that inputs a chest X-ray image and outputs the probability of pneumonia along with a heatmap. After training on an NIH dataset, the CheXNet network was tasked with a new set of 420 images. It outperformed four practicing Stanford University radiologists in diagnostic accuracy.

cal outcomes ranging from sepsis to hospital readmission. He found that sample sizes and accuracy varied widely, and all the reports were retrospective and not yet validated in real-world clinical settings. Real-world clinical validation, he argues, is important, because a model's accuracy doesn't guarantee it will work in a clinical setting and improve outcomes.

No clinician involved

That's a cue for an entrance by IDx, a company based in a city outside of Iowa City, Iowa. Topol says its technology was the first to be the subject of a prospective, peer-reviewed study of an AI application in a real-world setting.

IDx's technology uses machine learning to detect diabetic retinopathy (DR) without review from an eye specialist. Early detection of diabetic retinopathy is important because it can lead to treatment that prevents blindness, but half of patients with diabetes don't get annual eye exams. The IDx system has two algorithms. The

and IDx gets a cut of their reimbursement. In an email, IDx only verified that providers typically bill for the exams.

The company has touted results from a clinical trial that enrolled 900 people with diabetes. Findings from the IDx-funded study, reported in the August 2018 issue of *Nature Digital Medicine*, showed that, compared with the Wisconsin Fundus Photograph Reading Center—historically the gold standard for trials grading the severity of DR—IDx correctly identified the presence of more than mild DR 87% of the time (sensitivity), and correctly identified when patients didn't have the disease 90% of the time (specificity). Image quality was sufficient to render a diagnosis for 96.1% of participants.

University of Iowa (UI) Health Care, an integrated health system that includes the academic medical center where ophthalmology professor Michael Abramoff developed IDx before spinning it out into a company, was the first organization to adopt IDx-DR. The health system began using it at its diabetes clinic in June 2018. Approximately 100 patients have been screened so far. Brooks Jackson, MD, UI Health Care's vice president for medical affairs and dean of the medical school, says that "AI is very foreign to many physicians. They're still struggling with the electronic medical record."

Transplanting diabetic retinopathy testing from ophthalmology to primary care settings changes workflows, so careful planning is needed to minimize disruption, training, and added testing, says Jackson. Integrating IDx with the electronic medical record has required "some tweaking," he says.

Jackson believes the benefits, especially for patients, significantly outweigh any short-term inconvenience. Although the University of Iowa and IDx have a financial relationship—the university has a small stake in the company's patents and equity—Jackson says the opportunity to increase patient access to a screening that can help prevent blindness drove the decision to implement IDx-DR.

Because implementation began less than a year ago, IDx-DR's impact on UI's HEDIS metric requiring annual eye exams for patients with diabetes isn't yet known, but Jackson says things are going in the right direction. The health system eventually wants to get a measurement of the impact of IDx-DR on patient outcomes.

In addition to the clinical and patient access benefits, Jackson says that the technology has



With machine learning technology developed by a company called IDx, someone with just four hours of training can detect diabetic retinopathy in patients without having to have the diagnosis reviewed by an eye specialist. Real-world tested and FDA approved.

first detects whether the image quality is sufficient to make a diagnosis. The second provides a binary diagnosis and recommendation: either "more than mild DR—refer to eye care professional" or "negative—retest in 12 months." The company says that with four hours of training, someone with a high school diploma can operate the system. The FDA press release that announced the agency's approval of the system described it as the "first device authorized for marketing that provides a screening decision without the need for a clinician to also interpret the image or results."

IDx says it has more than 10 customers, including endocrinology and internal medicine clinics and academic medical centers. Multiple publications have reported that under IDx's revenue model, providers bill for the screening

generated positive profit margins for primary care practices while enabling eye specialists to focus on eyesight-saving procedures. UI Health Care will begin expanding IDx-DR to 35 primary care sites this summer, says Jackson.

GIGO still applies

IDx's results so far are promising, but AI is still largely unproven. Discerning whether an AI system is truly intelligent is going to be tricky. Kocher, the venture capitalist, says that intelligent technology requires data scientists who will thoughtfully assemble and review data to find errors, omissions, and biases. "We should be thinking about testing data sets like they're a drug," he says.

IBM Watson's foray into health care underscores the risks of a lax approach to data. Artificial intelligence does not suspend the rules of GIGO: garbage in, garbage out. Investigative reporting by *Stat* last year showed that the Watson computers were fed hypothetical rather than real patient data and that recommendations were based on specialists' expertise, not on evidence.

The FDA is tackling AI safety by subjecting AI-based software intended to "treat, diagnose, cure, mitigate, or prevent disease" to the same approval process that medical devices face, although there has been some criticism that the standards for device approvals are too low.

However, the FDA's requirement that a device must be reviewed each time it undergoes a major change could be problematic for AI-based products that continuously learn through new data, highlighting just one challenge in regulating this new category of technology. In April, the FDA sought public comment on a proposed regula-

tory framework that would permit algorithms to adapt without undergoing additional review.

Perpetuating a problem

Patient safety isn't the only worry. AI could perpetuate, even exacerbate, inequities in health care. Algorithms assume the data they're provided are reliable, so when trained with data sets that don't adequately sample a specific group, their "thinking" may be blind to issues particular to that group. In this way, AI is all-too human and flawed. Ethnic populations and women are often underrepresented in health care data sets.

Earlier this year, Kocher coauthored a blog post with Ezekiel Emanuel, the Obama administration health official who now chairs the University of Pennsylvania's department of medical ethics and health policy, that advocated for creating national test data sets with and without known biases to evaluate how effectively models can avoid perpetuating bias.

Even if algorithms ultimately prove accurate, safe, and unbiased, even the brainiest algorithm won't be able to conjure up the political will to deal with the health care system's excesses. In an opinion piece published in *JAMA* earlier this year, Emanuel and Robert Wachter, chair of the University of California–San Francisco Department of Medicine, warned that AI alone can't change patient and clinician behavior. "The most pressing problem with the U.S. health care system is not a lack of data or analytics but changing the behavior of millions of patients and clinicians," they said.

Or put simply, AI is ultimately a single tool, not a panacea, says Megan Zweig, a research leader at Rock Health.

—Sarah Kwon

How blockchain has strange bedfellows singing kumbaya

Competitors are coming together to see if they can leverage blockchain's magic to solve big problems. Provider directories are among the first use cases.

Mike Jacobs works in technology engineering at Optum, UnitedHealth Group's technology and health care services unit. Kyle Culver has a similar job at Humana. UnitedHealth and Humana are competitors, but now they're also collaborators,

in part because Jacobs and Culver first encountered each other at a workshop on blockchain, and then kept encountering each other on the health care tech circuit, speaking at conferences and holding workshops together.

Jacobs was working on a use case in which