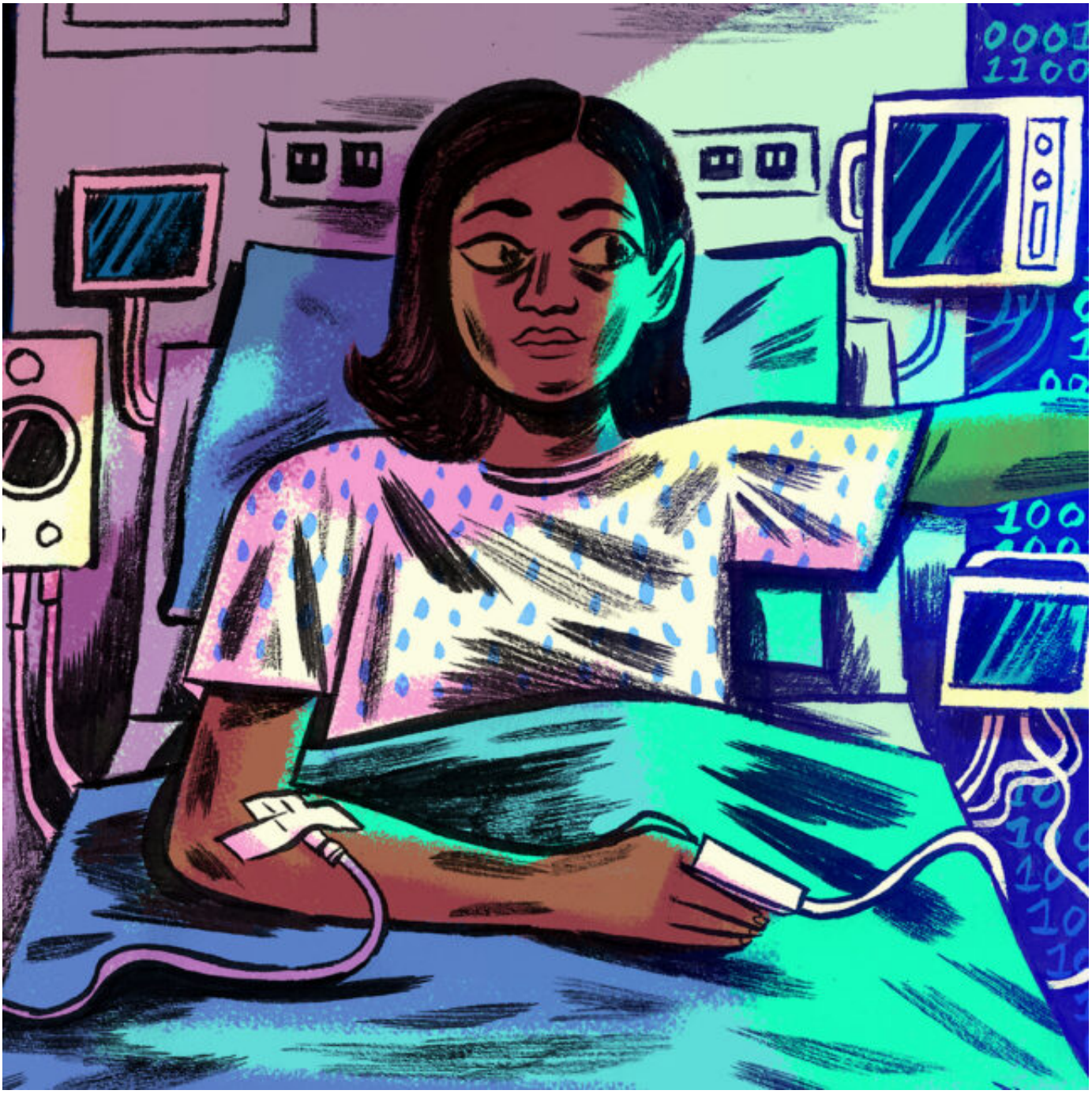
An invisible hand: Patients aren't being told about the AI systems advising their care

Rebecca Robbins



*Mike Reddy for STAT*none

Since February of last year, tens of thousands of patients hospitalized at one of Minnesota's largest health systems have had their discharge planning decisions informed with help from an artificial intelligence model. But few if any of those patients has any idea about the AI involved in their care.

That's because frontline clinicians at M Health Fairview generally don't mention the AI whirring behind the scenes in their conversations with patients.

At a growing number of prominent hospitals and clinics around the country, clinicians are turning to AI-powered decision support tools — many of them unproven — to help predict whether hospitalized patients are likely to develop complications or deteriorate, whether they're at risk of readmission, and whether they're likely to die soon. But these patients and their family members are often not informed about or asked to consent to the use of these tools in their care, a STAT examination has found.

The result: Machines that are completely invisible to patients are increasingly guiding decision-making in the clinic.

Hospitals and clinicians "are operating under the assumption that you do not disclose, and that's not really something that has been defended or really thought about," Harvard Law School professor Glenn Cohen said. Cohen is the author of one of only a few articles examining the issue, which has received surprisingly scant attention in the medical literature even as research about AI and machine learning proliferates.

In some cases, there's little room for harm: Patients may not need to know about an AI system that's nudging their doctor to move up an MRI scan by a day, like the one deployed by M Health Fairview, or to be more thoughtful, such as with algorithms meant to encourage clinicians to broach end-of-life conversations. But in other cases, lack of disclosure means that patients may never know what happened if an AI model makes a faulty recommendation that is part of the reason they are denied needed care or

undergo an unnecessary, costly, or even harmful intervention.

That's a real risk, because some of these AI models are fraught with bias, and even those that have been demonstrated to be accurate largely haven't yet been shown to improve patient outcomes. Some hospitals don't share data on how well the systems work, justifying the decision on the grounds that they are not conducting research. But that means that patients are not only being denied information about whether the tools are being used in their care, but also about whether the tools are actually helping them.

The decision not to mention these systems to patients is the product of an emerging consensus among doctors, hospital executives, developers, and system architects, who see little value — but plenty of downside — in raising the subject.

They worry that bringing up AI will derail clinicians' conversations with patients, diverting time and attention away from actionable steps that patients can take to improve their health and quality of life. Doctors also emphasize that they, not the AI, make the decisions about care. An AI system's recommendation, after all, is just one of many factors that clinicians take into account before making a decision about a patient's care, and it would be absurd to detail every single guideline, protocol, and data source that gets considered, they say.

Internist Karyn Baum, who's leading M Health Fairview's rollout of the tool, said she doesn't bring up the AI to her patients "in the same way that I wouldn't say that the X-ray has decided that you're ready to go home." She said she would never tell a fellow clinician not to mention the model to a patient, but in practice, her colleagues generally don't bring it up either.

Four of the health system's 13 hospitals have now rolled out the hospital discharge planning tool, which was developed by the Silicon Valley AI company Qventus. The model is designed to identify hospitalized patients who are likely to be clinically ready to go home soon and flag steps that

might be needed to make that happen, such as scheduling a necessary physical therapy appointment.

Clinicians consult the tool during their daily morning huddle, gathering around a computer to peer at a dashboard of hospitalized patients, estimated discharge dates, and barriers that could prevent that from occurring on schedule. A screenshot of the tool provided by Qventus lists a hypothetical 76-year-old patient, N. Griffin, who is scheduled to leave the hospital on a Tuesday — but the tool prompts clinicians to consider that he might be ready to go home Monday, if he can be squeezed in for an MRI scan by Saturday.

Baum said she sees the system as "a tool to help me make a better decision — just like a screening tool for sepsis, or a CT scan, or a lab value — but it's not going to take the place of that decision," she said. To her, it doesn't make sense to mention to patients. If she did, Baum said, she could end up in a lengthy discussion with patients curious about how the algorithm was created.

That could take valuable time away from the medical and logistical specifics that Baum prefers to spend time talking about with patients flagged by the Qventus tool. Among the questions she brings up with them: How are the patient's vital signs and lab test results looking? Does the patient have a ride home? How about a flight of stairs to climb when they get there, or a plan for getting help if they fall?

Some doctors worry that while well-intentioned, the decision to withhold mention of these AI systems could backfire.

"I think that patients will find out that we are using these approaches, in part because people are writing news stories like this one about the fact that people are using them," said Justin Sanders, a palliative care physician at Dana-Farber Cancer Institute and Brigham and Women's Hospital in Boston. "It has the potential to become an unnecessary distraction and undermine

trust in what we're trying to do in ways that are probably avoidable."

Patients themselves are typically excluded from the decision-making process about disclosure. STAT asked four patients who have been hospitalized with serious medical conditions — kidney disease, metastatic cancer, and sepsis — whether they'd want to be told if an AI-powered decision support tool were used in their care. They expressed a range of views: Three said they wouldn't want to know if their doctor was being advised by such a tool. But a fourth patient spoke out forcefully in favor of disclosure.

"This issue of transparency and upfront communication must be insisted upon by patients," said Paul Conway, a 55-year-old policy professional who has been on dialysis and received a kidney transplant, both consequences of managing kidney disease since he was a teenager.

The AI-powered decision support tools being introduced in clinical care are often novel and unproven — but does their rollout constitute research?

Many hospitals believe the answer is no, and they're using that distinction as justification for the decision not to inform patients about the use of these tools in their care. As some health systems see it, these algorithms are tools being deployed as part of routine clinical care to make hospitals more efficient. In their view, patients consent to the use of the algorithms by virtue of being admitted to the hospital.

At UCLA Health, for example, clinicians use a neural network to pinpoint primary care patients at risk of being hospitalized or frequently visiting the emergency room in the next year. Patients are not made aware of the tool because it is considered a part of the health system's quality improvement efforts, according to Mohammed Mahbouba, who spoke to STAT in February when he was UCLA Health's chief data officer. (He has since left the health system.)

“This is in the context of clinical operations,” Mahbouba said. “It’s not a research project.”

Oregon Health and Science University uses a regression-powered algorithm to monitor the majority of its adult hospital patients for signs of sepsis. The tool is not disclosed to patients because it is considered part of hospital operations.

“This is meant for operational care, it is not meant for research. So similar to how you’d have a patient aware of the fact that we’re collecting their vital sign information, it’s a part of clinical care. That’s why it’s considered appropriate,” said Abhijit Pandit, OHSU’s chief technology and data officer.

But there is no clear line that neatly separates medical research from hospital operations or quality control, said Pilar Ossorio, a professor of law and bioethics at the University of Wisconsin-Madison. And researchers and bioethicists often disagree on what constitutes one or the other.

“This has been a huge issue: Where is that line between quality control, operational control, and research? There’s no widespread agreement,” Ossorio said.

To be sure, there are plenty of contexts in which hospitals deploying AI-powered decision support tools are getting patients’ explicit consent to use them. Some do so in the context of clinical trials, while others ask permission as part of routine clinical operations.

At Parkland Hospital in Dallas, where the orthopedics department has a tool designed to predict whether a patient will die in the next 48 hours, clinicians inform patients about the tool and ask them to sign onto its use.

“Based on the agreement we have, we have to have patient consent explaining why we’re using this, how we’re using it, how we’ll use it to connect them to the right services, etc.,” said Vikas Chowdhry, the chief

analytics and information officer for a nonprofit innovation center incubated out of Parkland Health System in Dallas.

Hospitals often navigate those decisions internally, since manufacturers of AI systems sold to hospitals and clinics generally don't make recommendations to their customers about what, if anything, frontline clinicians should say to patients.

Jvion — a Georgia-based health care AI company that markets a tool that assesses readmission risk in hospitalized patients and suggests interventions to prevent another hospital stay — encourages the handful of hospitals deploying its model to exercise their own discretion about whether and how to discuss it with patients. But in practice, the AI system usually doesn't get brought up in these conversations, according to John Frownfelter, a physician who serves as Jvion's chief medical information officer.

"Since the judgment is left in the hands of the clinicians, it's almost irrelevant," Frownfelter said.

When patients are given an unproven drug, the protocol is straightforward: They must explicitly consent to enroll in a clinical study authorized by the Food and Drug Administration and monitored by an institutional review board. And a researcher must inform them about the potential risks and benefits of taking the medication.

That's not how it works with AI systems being used for decision support in the clinic. These tools aren't treatments or fully automated diagnostic tools. They also don't directly determine what kind of therapy a patient may receive — all of which would make them subject to more stringent regulatory oversight.

Developers of AI-powered decision support tools generally don't seek approval from the FDA, in part because the 21st Century Cures Act, which

was signed into law in 2016, was interpreted as taking most medical advisory tools out of the FDA's jurisdiction. (That could change: In guidelines released last fall, the agency said it intends to focus its oversight powers on AI decision-support products meant to guide treatment of serious or critical conditions, but whose rationale cannot be independently evaluated by doctors — a definition that lines up with many of the AI models that patients aren't being informed about.)

The result, for now, is that disclosure around AI-powered decision support tools falls into a regulatory gray zone — and that means the hospitals rolling them out often lack incentive to seek informed consent from patients.

"A lot of people justifiably think there are many quality-control activities that health care systems should be doing that involve gathering data," Wisconsin's Ossorio said. "And they say it would be burdensome and confusing to patients to get consent for every one of those activities that touch on their data."

In contrast to the AI-powered decision support tools, there are a few commonly used algorithms subject to the regulation laid out by the Cures Act, such as the type behind the genetic tests that clinicians use to chart a course of treatment for a cancer patient. But in those cases, the genetic test is extremely influential in determining what kind of therapy or drug a patient may receive. Conversely, there's no similarly clear link between an algorithm designed to predict whether a patient may be readmitted to the hospital and the way they'll be treated if and when that occurs.

"If it were me, I'd say just file for institutional review board approval and either get consent or justify why you could waive it."

Pilar Ossorio, professor of law and bioethics, University of Wisconsin-Madison

Still, Ossorio would support an ultra-cautious approach: "I do think people

throw a lot of things into the operations bucket, and if it were me, I'd say just file for institutional review board approval and either get consent or justify why you could waive it."

Further complicating matters is the lack of publicly disclosed data showing whether and how well some of the algorithms work, as well as their overall impact on patients. The public doesn't know whether OHSU's sepsis-prediction algorithm actually predicts sepsis, nor whether UCLA's admissions tool actually predicts admissions.

Some AI-powered decision support tools are supported by early data presented at conferences and published in journals, and several developers say they're in the process of sharing results: Jvion, for example, has submitted to a journal for publication a study that showed a 26% reduction in readmissions when its readmissions risk tool was deployed; that paper is currently in review, according to Jvion's Frownfelter.

But asked by STAT for data on their tools' impact on patient care, several hospital executives declined or said they hadn't completed their evaluations.

A spokesperson from UCLA said it had yet to complete an assessment of the performance of its admissions algorithm.

"Before you use a tool to do medical decision-making, you should do the research."

Pilar Ossorio, professor of law and bioethics, University of Wisconsin-Madison

A spokesperson from OHSU said that according to its latest report, run before the Covid-19 pandemic began in March, its sepsis algorithm had been used on 18,000 patients, of which it had flagged 1,659 patients as at-risk with nurses indicating concern for 210 of them. He added that the tool's impact on patients — as measured by hospital death rates and length of time

spent in the facility — was inconclusive.

“It’s disturbing that they’re deploying these tools without having the kind of information that they should have,” said Wisconsin’s Ossorio. “Before you use a tool to do medical decision-making, you should do the research.”

Ossorio said it may be the case that these tools are merely being used as an additional data point and not to make decisions. But if health systems don’t disclose data showing how the tools are being used, there’s no way to know how heavily clinicians may be leaning on them.

“They always say these tools are meant to be used in combination with clinical data and it’s up to the clinician to make the final decision. But what happens if we learn the algorithm is relied upon over and above all other kinds of information?” she said.

There are countless advocacy groups representing a wide range of patients, but no organization exists to speak for those who’ve unknowingly had AI systems involved in their care. They have no way, after all, of even identifying themselves as part of a common community.

STAT was unable to identify any patients who learned after the fact that their care had been guided by an undisclosed AI model, but asked several patients how they’d feel, hypothetically, about an AI system being used in their care without their knowledge.

Conway, the patient with kidney disease, maintained that he would want to know. He also dismissed the concern raised by some physicians that mentioning AI would derail a conversation. “Woe to the professional that as you introduce a topic, a patient might actually ask questions and you have to answer them,” he said.

Other patients, however, said that while they welcomed the use of AI and other innovations in their care, they wouldn’t expect or even want their

doctor to mention it. They likened it to not wanting to be privy to numbers around their prognosis, such as how much time they might expect to have left, or how many patients with their disease are still alive after five years.

“Any of those statistics or algorithms are not going to change how you confront your disease — so why burden yourself with them, is my philosophy,” said Stacy Hurt, a patient advocate from Pittsburgh who received a diagnosis of metastatic colorectal cancer in 2014, on her 44th birthday, when she was working as an executive at a pharmaceutical company. (She is now doing well and is approaching five years with no evidence of disease.)

Katy Grainger, who lost the lower half of both legs and seven fingertips to sepsis, said she would have supported her care team using an algorithm like OHSU’s sepsis model, so long as her clinicians didn’t rely on it too heavily. She said she also would not have wanted to be informed that the tool was being used.

“I don’t monitor how doctors do their jobs. I just trust that they’re doing it well.”

Katy Grainger, patient who developed sepsis

“I don’t monitor how doctors do their jobs. I just trust that they’re doing it well,” she said. “I have to believe that — I’m not a doctor and I can’t control what they do.”

Still, Grainger expressed some reservations about the tool, including the idea that it may have failed to identify her. At 52, Grainger was healthy and fairly young when she developed sepsis. She had been sick for days and visited an urgent care clinic, which gave her antibiotics for what they thought was a basic bacterial infection, but which quickly progressed to a serious case of sepsis.

“I would be worried that [the algorithm] could have missed me. I was young — well, 52 — healthy, in some of the best shape of my life, eating really well, and then boom,” Grainger said.

Dana Deighton, a marketing professional from Virginia, suspects that if an algorithm scanned her data back in 2013, it would have made a dire prediction about her life expectancy: She had just been diagnosed with metastatic esophageal cancer at age 43, after all. But she probably wouldn’t have wanted to hear about an AI’s forecast at such a tender and sensitive time.

“If a physician brought up AI when you are looking for a warmer, more personal touch, it might actually have the opposite and worse effect,” Deighton said. (She’s doing well now — her scans have turned up no evidence of disease since 2015.)

Harvard’s Cohen said he wants to see hospital systems, clinicians, and AI manufacturers come together for a thoughtful discussion around whether they should be disclosing the use of these tools to patients — “and if we’re not doing that, then the question is why aren’t we telling them about this when we tell them about a lot of other things,” he said.

Cohen said he worries that uptake and trust in AI and machine learning could plummet if patients “were to find out, after the fact, that there’s a rash of this being used without anyone ever telling them.”

“That’s a scary thing,” he said, “if you think this is the way the future is going to go.”

This is part of a yearlong series of articles exploring the use of artificial intelligence in health care that is partly funded by a grant from the Commonwealth Fund.