

TOM SIMONITE BUSINESS JUN 21, 2021 11:00 AM

An Algorithm That Predicts Deadly Infections Is Often Flawed

A study found that a system used to identify cases of sepsis missed most instances and frequently issued false alarms.



PHOTOGRAPH: JUAN BARRETO/GETTY IMAGES



The AI Database →

APPLICATION: PREDICTION END USER: SMALL COMPANY SECTOR: HEALTH CARE

TECHNOLOGY: MACHINE LEARNING

A COMPLICATION OF infection known as sepsis is the number one killer in US hospitals. So it's not surprising that more than 100 health systems use an early warning system offered by Epic Systems, the dominant provider of US electronic health records. The system throws up alerts based on a proprietary formula tirelessly watching for signs of the condition in a patient's test results.

But a new study using data from nearly 30,000 patients in University of Michigan hospitals suggests Epic's system performs poorly. The authors say it missed two-thirds of sepsis cases, rarely found cases medical staff did not notice, and frequently issued false alarms.

Karandeep Singh, an assistant professor at University of Michigan who led the study, says the findings illustrate a broader problem with the proprietary algorithms increasingly used in health care. "They're very widely used, and yet there's very little published on these models," Singh says. "To me that's shocking."

The study was published Monday in *JAMA Internal Medicine*. An Epic spokesperson disputed the study's conclusions, saying the company's system has "helped clinicians save thousands of lives."

Epic's is not the first widely used health algorithm to trigger concerns that technology supposed to improve health care is not delivering, or even actively harmful. In 2019, a system used on millions of patients to prioritize access to special care for people with complex needs was found to lowball the needs of Black patients compared to white patients. That prompted some Democratic senators to ask federal regulators to investigate bias in health algorithms. A study published in April found that statistical models used to predict suicide risk in mental health patients performed well for white and Asian patients but poorly for Black patients.

The models are “very widely used, and yet there’s very little published on them.”

— KARANDEEP SINGH, ASSISTANT PROFESSOR, UNIVERSITY OF MICHIGAN

The way sepsis stalks hospital wards has made it a special target of algorithmic aids for medical staff. Guidelines from the Centers for Disease Control and Prevention to health providers on sepsis encourage use of electronic medical records for surveillance and predictions. Epic has several competitors offering commercial warning systems, and some US research hospitals have built their own tools.

Automated sepsis warnings have huge potential, Singh says, because key symptoms of the condition, such as low blood pressure, can have other causes, making it difficult for staff to spot early. Starting sepsis treatment such as antibiotics just an hour sooner can make a big difference to patient survival. Hospital administrators often take special interest in sepsis response, in part because it contributes to US government hospital ratings.

Singh runs a lab at Michigan researching applications of machine learning to patient care. He got curious about Epic’s sepsis warning system after being asked to chair a committee at the university’s health system created to oversee uses of machine learning.

As Singh learned more about the tools in use at Michigan and other health systems, he became concerned that they mostly came from vendors that disclosed little about how they worked or performed. His own system had a license to use Epic’s sepsis prediction model, which the company told customers was highly accurate. But there had been no independent validation of its performance.

Singh and Michigan colleagues tested Epic’s prediction model on records for nearly 30,000 patients covering almost 40,000 hospitalizations in 2018 and 2019. The researchers noted how often Epic’s algorithm flagged people who developed sepsis as defined by the CDC and the Centers for Medicare and Medicaid Services. And they compared the alerts that the system would have triggered with sepsis treatments logged by staff, who did not see Epic sepsis alerts for patients included in the study.

The researchers say their results suggest Epic's system wouldn't make a hospital much better at catching sepsis and could burden staff with unnecessary alerts. The company's algorithm did not identify two-thirds of the roughly 2,500 sepsis cases in the Michigan data. It would have alerted for 183 patients who developed sepsis but had not been given timely treatment by staff.

At the same time, most of the Epic system's alerts would have been false alarms. When it flagged a patient, there was only a 12 percent chance that person would develop sepsis. "For all that alerting, you get very little value," Singh says. He believes the system could contribute to what people in health care call alert fatigue, the cavalcade of pop-ups, pings, and beeps that can cause physicians and nurses to feel overwhelmed and start ignoring notifications.

Keep Reading

Search our [artificial intelligence database](#) and discover stories by sector, tech, company, and more.

The Michigan authors say Epic tells customers its sepsis warning system can correctly distinguish two patients with and without sepsis at least 76 percent of the time. Their evaluation found it could do so only 63 percent of the time.

Singh says Epic's figures appear to make its system look more useful because they compare its alerts against records of billing codes for sepsis treatment. That effectively sets a lower bar for good performance, because it ignores sepsis cases not

detected by medical staff. “I think it’s developed to predict the wrong thing,” Singh says. “No one uses billing codes for detecting who has sepsis in a study.”

The Epic spokesperson pointed to a conference abstract published in January by Prisma Health of South Carolina on a smaller sample of 11,500 patients. It found that Epic’s system was associated with a 4 percent reduction in mortality of sepsis patients. Singh says that study used billing codes to define sepsis, not the clinical criteria medical researchers typically use.

Epic also says the Michigan study set a low threshold for sepsis alerts, which would be expected to produce a higher number of false positives; Singh says the threshold was chosen based on guidance from Epic.

See What’s Next in Tech With the Fast Forward Newsletter

From artificial intelligence and self-driving cars to transformed cities and new startups, sign up for the latest news.

By signing up you agree to our [User Agreement](#) (including the [class action waiver and arbitration provisions](#)), our [Privacy Policy & Cookie Statement](#) and to receive marketing and account-related emails from WIRED. You can unsubscribe at any time.

Roy Adams, an assistant professor who works on machine learning for health data at Johns Hopkins School of Medicine, wants to see other studies kick the tires on health algorithms shaping patient care. “We need more independent evaluations of these proprietary systems,” he says.

Adams says systems like Epic’s are becoming more common, but hospital administrators assessing them often have little data on how they operate, or perform in the clinic. Even where evaluation data is available, there aren’t clear standards on how to compare different systems.

Singh and other researchers are working on defining standardized ways to describe and compare the performance of health algorithms. He says Epic has recently made it easier for health care providers and other companies to integrate their own prediction models with the company’s record system, which should encourage more transparency and competition.

Singh also thinks that regulators should take more interest in systems like Epic's sepsis predictor. Recent guidance from the Food and Drug Administration about machine learning models in health care and interest in bias in machine learning from the White House Office of Science and Technology Policy make Singh feel optimistic that companies like Epic may soon have more incentive to be more rigorous and open with their algorithms.

More Great WIRED Stories

- 📧 The latest on tech, science, and more: [Get our newsletters!](#)
- One man's amazing journey to the [center of a bowling ball](#)
- The pandemic put an end to rush hour. [What happens now?](#)
- Want to write better? [Here are some tools to help](#)
- Facial verification [won't fight fraud](#)
- Watch drones fly through a fake forest [without crashing](#)
- 👁 Explore AI like never before with [our new database](#)
- 🎮 WIRED Games: Get the latest [tips, reviews, and more](#)
- 💻 Upgrade your work game with our Gear team's [favorite laptops](#), [keyboards](#), [typing alternatives](#), and [noise-canceling headphones](#)

