STAT+

How artificial intelligence could make clinical trials smarter

By [Katie Palmer](#) April 7, 2021

Adobe

Clinical trials have a dirty little secret. For all the careful work that goes into randomizing and blinding participants *just so*, the criteria that determine who can enter a trial can be unexpectedly arbitrary. Patients can be nixed because of age, lab values, medication history, and a laundry list of other factors that may not always be necessary.

"It was certainly surprising to us that these clinical trial criteria designs are fairly ad hoc and quite anecdotal," said James Zou, who leads Stanford's Laboratory for Machine Learning, Genomics, and Health. The unwitting result can be that

women, older patients, and people of color are excluded from studies at higher rates. That's why Zou, in collaboration with Genentech and colleagues at Stanford, started looking at how to design smarter eligibility criteria that can boost enrollment without compromising on safety.

In a paper published Wednesday in [Nature](#), they describe an artificial intelligence tool to run simulated clinical trials for cancer drugs applying different eligibility criteria. Using data from real-life patients, they found that in most cases, they could loosen the criteria for trial entry — making it possible to include more and more diverse participants — without any impact on safety.

That can benefit patients immensely. "Without the real-world data, you never know what will happen to patients who are filtered out by the previous studies," said Ruishan Liu, a Ph.D. candidate in Zou's lab and lead author on the paper.

But in their simulated trials, which were based on real-world studies for non-small cell lung cancer, they saw evidence that these previously excluded participants would benefit just as much from these trials — and sometimes more — than the original group.

It also benefits pharma companies, which could double the size of their eligible patient populations, according to the analysis. That would make it easier, faster, and cheaper to complete trials. And they could stand to see even stronger results for their drugs: "If I'm the company, I want to get the biggest signal, and oftentimes a lot of people I'm excluding are the ones that actually give me the bigger, more striking signal on my drugs," said Zou.

Genentech is building an internal version of the tool, called Trial Pathfinder, to aid in the design of oncology clinical trials.

Tools such as Trial Pathfinder could present a more streamlined approach to the design of inclusion and exclusion criteria, which can be slapdash at times. In some cases, researchers will just copy and paste the criteria from previous successful trials, said Chunhua Weng, a biomedical informatician at Columbia

and author of a perspective published with the study, both because it's easy and because it's a conservative way to reduce risk.

"Very often, no one knows why a certain eligibility criterion goes into a particular study," said Weng. "They just feel like, 'OK, it's already there in the literature.'"

But the availability of real-world data from electronic health records has started to change that. In 2018, the Food and Drug Administration put together a framework for the use of [real-world evidence](#) to support its regulatory approach. "That actually was a watershed, because prior to that there really was no application of real-world evidence for any regulatory decisions," said Jeff Elton, CEO of ConcertAI, which supports clinical trials for oncology companies.

The Stanford and Genentech group used highly structured EHR data from Flatiron Health, a startup [acquired by Swiss drug company Roche Holding](#) in 2018. Flatiron spent years assembling its data on cancer patients in community settings, labeling it carefully to enable easier analysis. "Now the datasets are tens of thousands of patients, and that's a significant portion of the treated population in that particular category," said Elton.

But most real-world datasets are messy. Not every disease has such clean patient records to work with, or a parent company to acquire them. (Roche owns Genentech.)

"In reality, the EHR data will not be as clean as the Flatiron database," said Weng.

To develop evidence-based eligibility criteria for non-cancer trials, the data will need to catch up. Which diseases could be up next? "That's the $3 billion question," said Elton. "Oncology has been benefited by a huge amount of research and large scale datasets." But he guesses that immunology, rheumatology, and neurology could justify the type of investment Flatiron and ConcertAI have made in disease-specific real-world datasets.

To extend the work, researchers will have to optimize for other variables, too. "In this particular paper they only focus on the hazard ratio for survival," said Weng, which reflects the rate of death in the treated group compared to the control. "But

when it comes to other trials, the outcomes can be very complex." Say you're studying a drug for congenital blindness — the outcome to manage for isn't survival, but progression to sightlessness.

As Elton noted, pharma companies may also have nonclinical goals to reach, "such as increasing diversity." He points to the example of prostate cancer multiple myeloma, which has a disproportionately negative effect on certain Black male populations. "If I want to assure that I actually accrue to those populations, I may need to be very mindful of the design."

But that approach still requires siloed, disease-specific databases, which require top-down investment. Another strategy, Weng said, could be to go bottom-up.

She's a member of the Observational Health Data Sciences and Informatics community, which adopted a common data model called OMOP, for Observational Medical Outcomes Partnership, to standardize EHR data. With [hundreds of millions](#) of patient records formatted, the consortium has enabled large-scale observational studies across the globe.

"If pharma companies could pay attention to OMOP, all future data resources could be formatted using this data model," said Weng. "We will be on the right track to try to replicate the study reported by this Stanford group" — and give patients the inclusive, rational trial design that they deserve.

About the Author

![Katie Palmer portrait]

[Katie Palmer](#)

Health Tech Correspondent

Katie Palmer is a health tech correspondent at STAT.

katie.palmer@statnews.com
[@KatieMPalmer](#)