Key takeaways

Researchers today can use arti cial intelligence and data analytics to speed up the clinical trial process. Instead of having to recruit patients for a traditional control arm—the group that doesn't get the experimental treatment given to the test group—investigators are building "external control arms," which reuse data on control-group patients from past clinical trials.

External control arms yield several bene ts. They can reduce or eliminate the time normally needed to recruit control patients, expediting access to experimental treatment for patients in the test group. They cut the costs of recruiting control group patients and tracking them during the trial. And using external controls makes it easier to recruit potential participants, because everyone recruited will get the treatment.

The US Food and Drug Administration looks favorably on external control arms in general, especially in single-arm trials (a type of trial in which a regular control group is impractical). Replacing traditional control arms with external data faces more scrutiny—but a hybrid design, in which external controls supplement a recruited control arm, is currently under review by the FDA.

Clinical trials have never been more in the public eye than in the past year, as the world watched the development of vaccines against covid-19, the disease at the center of the 2020 coronavirus pandemic. Discussions of study phases, e cacy, and side e ects dominated the news. The most distinctive feature of the vaccine trials was their speed. Because the vaccines are meant for universal distribution, the study population is, basically, everyone. That unique feature means that recruiting enough people for the trials has not been the obstacle that it commonly is.

"One of the most di cult parts of my job is enrolling patients into studies," says Nicholas Borys, chief medical o cer for Lawrenceville, N.J., biotechnology company Celsion, which develops next-generation chemotherapy and immunotherapy agents for liver and ovarian cancers and certain types of brain tumors. Borys estimates that fewer than 10% of cancer patients are enrolled in clinical trials. "If we could get that up to 20% or 30%, we probably could have had several cancers conquered by now."

Clinical trials test new drugs, devices, and procedures to determine whether they're safe and e ective before they're approved for general use. But the path from study design to approval is long, winding, and expensive. Today, researchers are using arti cial intelligence and advanced data analytics to speed up the process, reduce costs, and get e ective treatments more swiftly to those who need them. And they're tapping into an underused but rapidly growing resource: data on patients from past trials.

Building external controls

Clinical trials usually involve at least two groups, or "arms": a test or experimental arm that receives the treatment under investigation, and a control arm that doesn't. A control arm may receive no treatment at all, a placebo or the current standard of care for the disease being treated, depending on what type of treatment is being studied and what it's being compared with under the study protocol.

It's easy to see the recruitment problem for investigators studying therapies for cancer and other deadly diseases: patients with a life-threatening condition need help now. While they might be willing to take a risk on a new treatment, "the last thing they want is to be randomized to a control arm," Borys says. Combine that reluctance with the need to recruit patients who have relatively rare diseases—for example, a form of breast cancer characterized by a speci c genetic marker—and the time to recruit enough people can stretch out for months, or even years. Nine out of 10 clinical trials worldwide—not

External control arms can make clinical trials less expensive by reducing the number of patients that need active management.

just for cancer but for all types of conditions—can't recruit enough people within their target timeframes. Some trials fail altogether for lack of enough participants.

What if researchers didn't need to recruit a control group at all and could o er the experimental treatment to everyone who agreed to be in the study? Celsion is exploring such an approach with New York-headquartered Medidata, which provides management software and electronic data capture for more than half of the world's clinical trials, serving most major pharmaceutical and medical device companies, as well as academic medical centers. Acquired by French software company Dassault Systèmes in 2019, Medidata has compiled an enormous "big data" resource: detailed information from more than 23,000 trials and nearly 7 million patients going back about 10 years.

The idea is to reuse data from patients in past trials to create "external control arms." These groups serve the same function as traditional control arms, but they can be used in settings where a control group is di cult to recruit: for extremely rare diseases, for example, or conditions such as cancer, which are imminently life-threatening. They can also be used e ectively for "single-arm" trials, which make a control group impractical: for example, to

measure the e ectiveness of an implanted device or a surgical procedure. Perhaps their most valuable immediate use is for doing rapid preliminary trials, to evaluate whether a treatment is worth pursuing to the point of a full clinical trial.

Medidata uses arti cial intelligence to plumb its database and nd patients who served as controls in past trials of treatments for a certain condition to create its proprietary version of external control arms. “We can carefully sel Txternal contan cefullmspid prord [(Medida9 (f tr)20 (ea)10 (tmen)9.9 “)-i 0 -1.44s9.9 (mion o)10 (fch implcurv)10 (ers

Real-world data—the information about all of us that's in our electronic health records, our pharmacy's prescription database, our insurance claims—has the potential to further re ne the clinical trial process, make it faster, cheaper, and more accurate, and extend it to post-market surveillance to verify whether new medications and procedures ful ll the promise of their clinical trials.

Medidata is exploring such possibilities with San Francisco health tech company Datavant, which works with health systems and electronic-claims clearinghouses to collect and link data on patients—while preserving their privacy as required by federal law—so that researchers can study the health information that's gathered over time on a single individual, or do advanced analytics on a group of patients that share certain characteristics.

"We're trying to link data from those clinical trial cohorts to the rest of their real-world data, but in a privacy-preserving way," says Jason LaBonte, chief strategy o cer at Datavant. "So, if you want to understand more about the patients that were in the trial, you're not stuck if you didn't collect the data."

Datavant works with about 400 health systems and other providers and has access to a vast trove of insurance claim data. LaBonte estimates that among its various data sources, the company has at least some health and other information on about 300 million people in the United States.

Health systems, companies, and organizations that work with Datavant run its software on their patient databases to create a de-identi ed version of the data. The set of data that identi es someone as a unique individual is replaced with a "linking token." Each institution's data is encrypted so that no other institution can identify individual patients.

Datavant's secret sauce is a complex method of identifying the tokens for the same patient from multiple sources. In that way, a researcher doesn't know the identity of a subject but does know all the subject's diagnoses and treatments over time: across multiple physicians, hospitals, pharmacies, labs, and even insurers. The researcher can track what happens to an individual clinical trial participant after the trial is over—a type of long-term follow-up that's typically expensive and not always practical.

"If I have a subject in a clinical trial, and I want to connect all of her lab records and her electronic medical record, and her insurance claims, data linking via tokenization allows us to link all that data," says Arnaub Chatterjee of Medidata Acorn AI. "We can pull together all those records that exist out in the ether." Clinical trial patients need to give explicit consent for their data to be used in this manner.

"Everybody who uses our software retains full control of their data and can say no to anything they don't want to participate in, but the data is safe to share under HIPAA," LaBonte says, referring to the US health information privacy law. The sheer size of the database allows researchers to assemble a study population that includes almost any set of characteristics.

LaBonte predicts that real-world data will be used increasingly for "pragmatic trials" such as the studies done in 2020 establishing that the malaria drug hydroxychloroquine was not e ective against covid-19. "That was a good pragmatic trial candidate because doctors were using it in practice, so people could go through their data, nd the patients where it was tried, and then nd a matching set of patients who weren't given hydroxychloroquine and call it the control arm," LaBonte

countries. “You have to have a very large treatment e ect for this kind of crude comparison to be convincing, and

“Clinical trials are better, faster, cheaper with big data” is an executive brie nga5.5 10 (Cl Sy o)20 enology Reviewtarr© CIt et

Illustrations

Person icons: kornn. Circular puzzle piece: Arcady. Chart icon: Blan-k. Cardiovascular icon: Stmool. Server icon: Mark stock. Stopwatch icon: NikWB.

All illustrations assembled by Scott Shultz Design, sourced from Shutterstock.