Cutting to the chase

From study design to patient recruitment, researchers are investigating how AI could speed up clinical trials.
By Matthew Hutson

For decades, computing power followed Moore's law, advancing at a predictable pace. The number of components on an integrated circuit doubled roughly every two years. In 2012, researchers coined the term Eroom's law (Moore spelled backwards) to describe the contrasting path of drug development[1]. Over the previous 60 years, the number of drugs approved in the United States per billion dollars in R&D spending had halved every nine years. It can now take more than a billion dollars in funding and a decade of work to bring one new medication to market. Half of that time and money is spent on clinical trials, which are growing larger and more complex. And only one in seven drugs that enters phase I trials is eventually approved.

Some researchers are hoping that the fruits of Moore's law can help to curtail Eroom's law. Artificial intelligence (AI) has already been used to make strong inroads into the early stages of drug discovery, assisting in the search for suitable disease targets and new molecule designs. Now scientists are starting to use AI to manage clinical trials, including the tasks of writing protocols, recruiting patients and analysing data.

Reforming clinical research is "a big topic of interest in the industry", says Lisa Moneymaker, the chief technology officer and chief product officer at Saama, a software company in Campbell, California, that uses AI to help organizations automate parts of clinical trials. "In terms of applications," she says, "it's like a kid in a candy store."

Trial by design

The first step of the clinical-trials process is trial design. What dosages of drugs should be given? To how many patients? What data should be collected on them? The lab of Jimeng Sun, a computer scientist at the University of Illinois Urbana-Champaign, developed an algorithm called HINT (hierarchical interaction network) that can predict whether a trial will succeed, based on the drug molecule, target disease and patient eligibility criteria. They followed up with a system called SPOT (sequential predictive modelling of clinical trial outcome) that additionally takes into account when the trials in its training data took place and weighs more recent trials more heavily. Based on the predicted outcome, pharmaceutical companies might decide to alter a trial design, or try a different drug completely.

A company called Intelligent Medical Objects in Rosemont, Illinois, has developed SEETrials, a method for prompting OpenAI's large language model GPT-4 to extract safety and efficacy information from the abstracts of clinical trials. This enables trial designers to quickly see how other researchers have designed trials and what the outcomes have been. The lab of Michael Snyder, a geneticist at Stanford University in California, developed a tool last year called CliniDigest that simultaneously summarizes dozens of records from ClinicalTrials.gov, the main US registry for medical trials, adding references to the unified summary. They've used it to summarize how clinical researchers are using wearables such as smartwatches, sleep trackers and glucose monitors to gather patient data. "I've had conversations with plenty of practitioners who see wearables' potential in trials, but do not know how to use them for highest impact," says Alexander Rosenberg Johansen, a computer-science student in Snyder's lab. "Best practice does not exist yet, as the field is moving so fast."

Most eligible

The most time-consuming part of a clinical trial is recruiting patients, taking up to one-third of the study length. One in five trials don't even recruit the required number of people, and nearly all trials exceed the expected recruitment timelines. Some researchers would like to accelerate the process by relaxing some of the eligibility criteria while maintaining safety. A group at Stanford led by James Zou, a biomedical data scientist, developed a system called Trial Pathfinder that analyses a set of completed clinical trials and assesses how adjusting the criteria for participation – such as thresholds for blood pressure and lymphocyte counts – affects hazard ratios, or rates of negative incidents such as serious illness or death among patients. In one study[2], they applied it to drug trials for a type of lung cancer. They found

that adjusting the criteria as suggested by Trial Pathfinder would have doubled the number of eligible patients without increasing the hazard ratio. The study showed that the system also worked for other types of cancer and actually reduced harmful outcomes because it made sicker people – who had more to gain from the drugs – eligible for treatment.

AI can eliminate some of the guesswork and manual labour from optimizing eligibility criteria. Zou says that sometimes even teams working at the same company and studying the same disease can come up with different criteria for a trial. But now several firms, including Roche, Genentech and AstraZeneca, are using Trial Pathfinder. More recent work from Sun's lab in Illinois has produced AutoTrial, a method for training a large language model so that a user can provide a trial description and ask it to generate an appropriate criterion range for, say, body mass index.

Once researchers have settled on eligibility criteria, they must find eligible patients. The lab of Chunhua Weng, a biomedical informatician at Columbia University in New York City (who has also worked on optimizing eligibility criteria), has developed Criteria2Query. Through a web-based interface, users can type inclusion and exclusion criteria in natural language, or enter a trial's identification number, and the program turns the eligibility criteria into a formal database query to find matching candidates in patient databases.

Weng has also developed methods to help patients look for trials. One system, called DQueST, has two parts. The first uses Criteria2Query to extract criteria from trial descriptions. The second part generates relevant questions for patients to help narrow down their search. Another system, TrialGPT, from Sun's lab in collaboration with the US National Institutes of Health, is a method for prompting a large language model to find appropriate trials for a patient. Given a description of a patient and clinical trial, it first decides whether the patient fits each criterion in a trial and offers an explanation. It then aggregates these assessments into a trial-level score. It does this for many trials and ranks them for the patient.

Helping researchers and patients find each other doesn't just speed up clinical research. It also makes it more robust. Often trials unnecessarily exclude populations such as children, the elderly or people who are pregnant, but AI can find ways to include them. People with terminal cancer and those with rare diseases have an especially hard time finding trials to join. "These patients sometimes do more work than clinicians in diligently searching for trial opportunities," Weng says. AI can help match them with relevant projects.

AI can also reduce the number of patients needed for a trial. A start-up called Unlearn in San Francisco, California, creates digital twins of patients in clinical trials. Based on an experimental patient's data at the start of a trial, researchers can use the twin to predict how the same patient would have progressed in the control group and compare outcomes. This method typically reduces the number of control patients needed by between 20% and 50%, says Charles Fisher, Unlearn's founder and chief executive. The company works with a number of small and large pharmaceutical companies. Fisher says digital twins benefit not only researchers, but also patients who enrol in trials, because they have a lower chance of receiving the placebo.

Patient maintenance
The hurdles in clinical trials don't end once patients enrol. Drop-out rates are high. In one analysis of 95 clinical trials, nearly 40% of patients stopped taking the prescribed medication in the first year. In a recent review article[3], researchers at Novartis mentioned ways that AI can help. These include using past data to predict who is most likely to drop out so that clinicians can intervene, or using AI to analyse videos of patients taking their medication to ensure that doses are not missed.

> **"AI can find ways to include populations usually unnecessarily excluded from clinical trials."**

Chatbots can answer patients' questions, whether during a study or in normal clinical practice. One study[4] took questions and answers from Reddit's AskDocs forum and gave the questions to ChatGPT. Health-care professionals preferred ChatGPT's answers to the doctors' answers nearly 80% of the time. In another study[5], researchers created a tool called ChatDoctor by fine-tuning a large language model (Meta's LLaMA-7B) on patient-doctor dialogues and giving it real-time access to online sources. ChatDoctor could answer questions about medical information that was more recent than ChatGPT's training data.

Putting it together
AI can help researchers manage incoming clinical-trial data. The Novartis researchers reported that it can extract data from unstructured reports, as well as annotate images or lab results, add missing data points (by predicting values in results) and identify subgroups among a population that responds uniquely to

a treatment. Zou's group at Stanford has developed PLIP, an AI-powered search engine that lets users find relevant text or images within large medical documents. Zou says they've been talking with pharmaceutical companies that want to use it to organize all of the data that comes in from clinical trials, including notes and pathology photos. A patient's data might exist in different formats, scattered across different databases. Zou says they've also done work with insurance companies, developing a language model to extract billing codes from medical records, and that such techniques could also extract important clinical trial data from reports such as recovery outcomes, symptoms, side effects and adverse incidents.

To collect data for a trial, researchers sometimes have to produce more than 50 case report forms. A company in China called Taimei Technology is using AI to generate these automatically based on a trial's protocol.

A few companies are developing platforms that integrate many of these AI approaches into one system. Xiaoyan Wang, who heads the life-science department at Intelligent Medical Objects, co-developed AutoCriteria, a method for prompting a large language model to extract eligibility requirements from clinical trial descriptions and format them into a table. This informs other AI modules in their software suite, such as those that find ideal trial sites, optimize eligibility criteria and predict trial outcomes. Soon, Wang says, the company will offer ChatTrial, a chatbot that lets researchers ask about trials in the system's database, or what would happen if a hypothetical trial were adjusted in a certain way.

The company also helps pharmaceutical firms to prepare clinical-trial reports for submission to the US Food and Drug Administration (FDA), the organization that gives final approval for a drug's use in the United States. What the company calls its Intelligent Systematic Literature Review extracts data from comparison trials. Another tool searches social media for what people are saying about diseases and drugs in order to demonstrate unmet needs in communities, especially those that feel underserved. Researchers can add this information to reports.

Zifeng Wang, a student in Sun's lab, in Illinois, says he's raising money with Sun and another co-founder, Benjamin Danek, for a start-up called Keiji AI. A product called TrialMind will offer a chatbot to answer questions about trial design, similar to Xiaoyan Wang's. It will do things that might normally require a team of data scientists, such as write code to analyse data or produce visualizations. "There are a lot of opportunities" for AI in clinical trials, he says, "especially with the recent rise of larger language models."

At the start of the pandemic, Saama worked with Pfizer on its COVID-19 vaccine trial. Using Saama's AI-enabled technology, SDQ, they 'cleaned' data from more than 30,000 patients in a short time span. "It was the perfect use case to really push forward what AI could bring to the space," Moneymaker says. The tool flags anomalous or duplicate data, using several kinds of machine-learning approaches. Whereas experts might need two months to manually discover any issues with a data set, such software can do it in less than two days.

Other tools developed by Saama can predict when trials will hit certain milestones or lower drop-out rates by predicting which patients will need a nudge. Its tools can also combine all the data from a patient — such as lab tests, stats from wearable devices and notes — to assess outcomes. "The complexity of the picture of an individual patient has become so huge that it's really not possible to analyse by hand any-more," Moneymaker says.

Xiaoyan Wang notes that there are several ethical and practical challenges to AI's deploy-ment in clinical trials. AI models can be biased. Their results can be hard to reproduce. They require large amounts of training data, which could violate patient privacy or create security risks. Researchers might become too depend-ent on AI. Algorithms can be too complex to understand. "This lack of transparency can be problematic in clinical trials, where under-standing how decisions are made is crucial for trust and validation," she says. A recent review article[6] in the *International Journal of Surgery* states that using AI systems in clinical trials "can't take into account human faculties like common sense, intuition and medical training".

Moneymaker says the processes for designing and running clinical trials have often been slow to change, but adds that the FDA has relaxed some of its regulations in the past few years, leading to "a spike of innovation": decentralized trials and remote monitoring have increased as a result of the pandemic, opening the door for new types of data. That has coincided with an explosion of generative-AI capabilities. "I think we have not even scratched the surface of where generative-AI applicability is going to take us," she says. "There are problems we couldn't solve three months ago that we can solve now."

Matthew Hutson is a science writer based in New York City.

1. Scannell, J. W., Blanckley, A., Boldon, H. & Warrington, B. *Nature Rev. Drug. Discov.* **11**, 191–200 (2012).
2. Liu, R. *et al. Nature* **592**, 629–633 (2021).
3. Blaschke, T. F., Osterberg, L., Vrijens, B. & Urquhart, J. *Annu. Rev. Pharmacol. Toxicol.* **52**, 275–301 (2012).
4. Ayers, J. W. *et al. JAMA Intern. Med.* **183**, 589–596 (2023).
5. Li, Y. *et al. Cureus* **15**, e40895 (2023).
6. Chopra, H. *et al. Int. J. Surg.* **109**, 4211–4220 (2023).

GLOBAL PIPELINE

More than 7,000 drugs made it from phase I clinical trials to regulatory submission in 2022, almost 3,000 more than in 2007. The United States' slice of this global R&D pipeline has remained relatively stable, at above 40% for the past 15 years. Companies headquartered in China, meanwhile, have rapidly increased their share, from 2% in 2007 to 15% in 2022, according to analysis by IQVIA, a biotechnology company based in Durham, North Carolina.
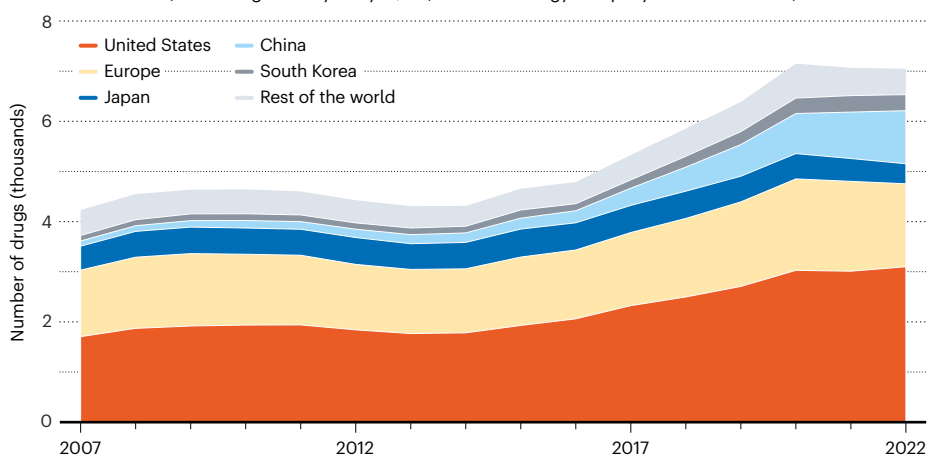
RECRUITMENT DRIVE

Clinical trials in infectious diseases have seen the biggest increase in participants in the past few years, with unprecedented numbers signing up because of the COVID-19 pandemic. But even when excluding COVID-19 and Ebola, trials for other infectious diseases had nearly 600,000 subjects in 2022, compared with 125,000 in 2018.
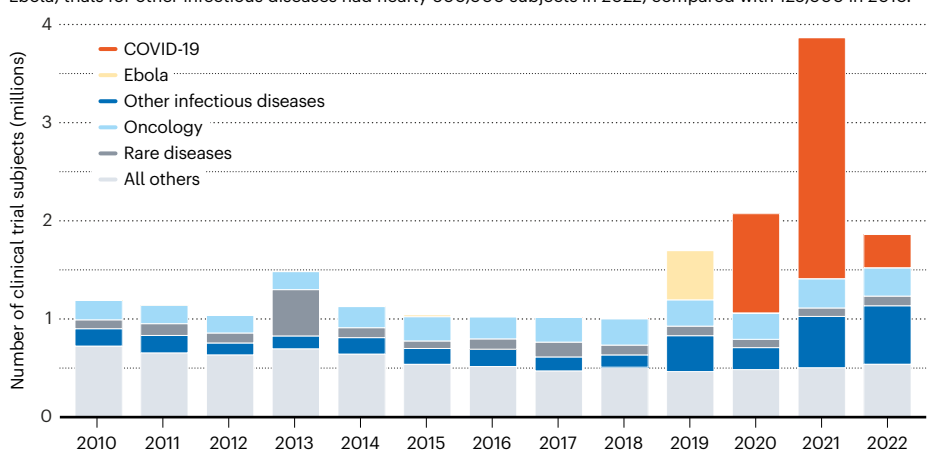
INCREASED TRIAL COMPLEXITY

The cost of drug development continues to rise, and the size and complexity of clinical trials is a major factor. In the past two decades, the number of countries in which a clinical trial is conducted has more than doubled, and the average number of data points collected has grown dramatically. There are more endpoints — outcomes of a clinical trial that help to determine the efficacy and safety of an experimental therapy — and procedures to measure these outcomes, such as blood tests and heart-activity assessments. By comparison, eligibility criteria for participants, which include demographics such as age and sex and whether a participant is a healthy or a patient volunteer, have remained relatively consistent.

Category	2001–05	2011–15	2016–20	20-year overall rise
Endpoints*	7	13	22	214%
Procedures*	110	187	263	139%
Eligibility criteria*	31	30	30	−3%
Countries*	6	9	15	150%
Data points collected*	494,236	929,203	3,453,133	599%

*Mean of total numbers