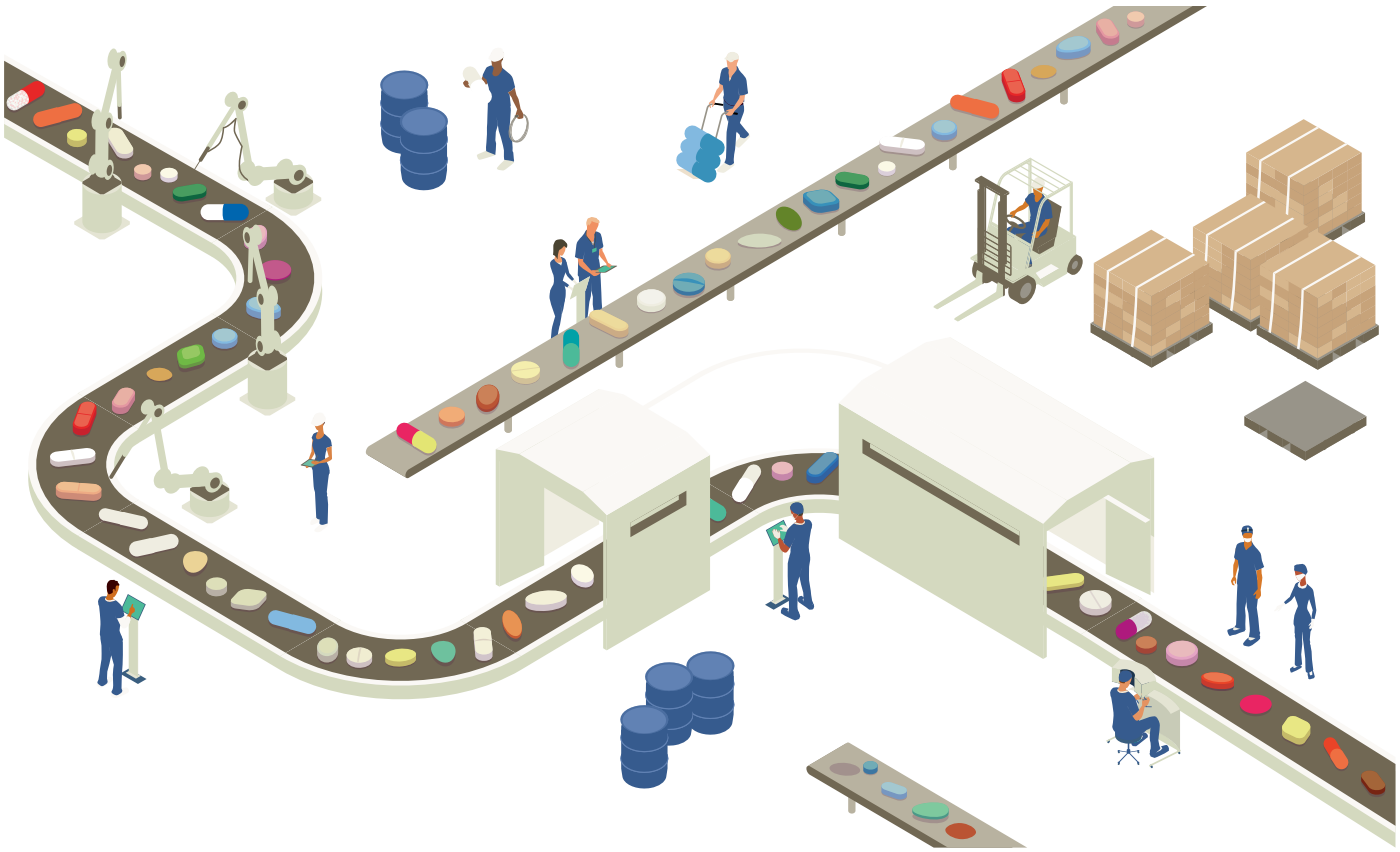
https://doi.org/10.1038/s41591-023-02361-0



INSIDE THE NASCENT INDUSTRY OF AI-DESIGNED DRUGS

Artificial intelligence tools are beginning to upend the drug discovery pipeline, with several new compounds entering clinical trials. **By Carrie Arnold**

Drug discovery is expensive, inefficient, and fraught with failure. An estimated 86% of drug candidates developed between 2000 and 2015 did not meet their stated endpoints.

Despite this challenge, the use of artificial intelligence (AI) and machine learning to understand drug targets better and synthesize chemical compounds to interact with them has not been easy to sell. Alex Zhavoronkov would know. When the CEO and founder of Insilico Medicine, with offices in Hong Kong and New York, first started trying to raise funding nearly a decade ago, he struggled to find others who shared his vision.

"It was such a grand goal, but every time I went to a venture capitalist, they never gave me money," says Zhavoronkov.

Even as recently as 5 years ago, his presentations had to explain to pharma collaborators why AI was so promising. Not anymore. Now he is at the forefront of drug discovery's AI nascent revolution.

"We've managed to get here in three years, and we didn't fail. And we did it multiple times," Zhavoronkov says.

The persistence of Zhavoronkov and a small cadre of other startup founders, including Exscientia's Andrew Hopkins and BenevolentAI's Bryn Williams-Jones, means that not only are some of the biggest players in pharma already convinced of the utility of AI in drug development, but also some of these drugs are beginning their ultimate test in clinical trials (Table 1).

"In the last couple of years, AI has gone from being hypothetically interesting to real programs moving towards the clinic," says Williams-Jones. "There's no shortcuts to drug discovery. We can have better informed ideas, but you still have to go through the rest of the [development] process."

These trials are still in their early days, says Hopkins, so it is not yet clear which compound will cross the finish line first. But he is

Table 1 | Selected AI-designed drugs in or entering clinical trials

Treatment	Organization	Description	Phase	Lead indication
REC-2282	Recursion	Small molecule pan-HDAC inhibitor	2/3	Neurofibromatosis type 2
REC-994	Recursion	Small molecule superoxide scavenger	2	Cerebral cavernous malformation
REC-4881	Recursion	Small molecule inhibitor of MEK1 and MEK2	2	Familial adenomatous polyposis
INS018_055	InSilico Medicine	Small molecule inhibitor	2	Idiopathic pulmonary fibrosis
BEN-2293	BenevolentAI	Topical pan-tyrosine kinase inhibitor	2a	Atopic dermatitis
EXS-21546	Exscientia and Evotec	A_{2A} receptor antagonist	1b/2	Solid tumors carrying high adenosine signatures.
RLY-4008	Relay Therapeutics	Inhibitor of FGFR2	1/2	FGFR2-altered cholangiocarcinoma
EXS-4318	Exscientia	PKC-θ inhibitor	1/2	Inflammatory and autoimmune conditions
BEN-8744	BenevolentAI	Small molecule PDE10 inhibitor	1	Ulcerative colitis
Undisclosed	Recursion	Small molecular inhibitor of RBM39, a CDK12-associated protein	Pre-clinical	HRD-negative ovarian cancer

confident that the use of AI is leaving an indelible mark on drug development and promises to make the process better, faster, and cheaper, as well as enabling the development of more first-in-class compounds.

"We expect this year to see some major advances in the number of molecules and approved drugs produced by generative AI methods that are moving forward," Hopkins says.

Entering trials

As AI-designed drugs enter clinical trials, pharma companies can see how their new compounds are paying off. The preliminary read-outs look promising. In June 2022, Exscientia announced preliminary results from a phase 1 trial of EXS-21546, a highly selective A_{2A} receptor antagonist developed with Germany's Hamburg-based Evotec. The small molecule has subsequently entered phase 1b/2 trials for patients with solid tumors carrying high adenosine signatures.

Exscientia's next AI-developed candidate, a small molecule called EXS4318, is not far behind. A selective protein kinase C-theta (PKC-θ) inhibitor, designed for inflammatory and autoimmune conditions, EXS4318 has been licensed to Bristol Myers Squibb in a partnership worth up to US$1.2 billion, according to a company press release. The company has 16 other AI-designed drugs in its pipeline, including drugs for COVID-19, tuberculosis, malaria, and hypophosphatasia – a rare, inherited disorder that affects bones and teeth.

"It's not just about using generative AI to help us to precision design an exact molecule," Hopkins says, "but also actually helping us precision design which patients are responders and non-responders."

What this would look like in practice, Hopkins says, is performing deep, multi-omics (single-cell proteomics, transcriptomics, and genomics) analyses of participants before the trial starts to identify multi-gene signature biomarkers. This will help the researchers to determine which participants are most likely to respond – and why. At the end of the trial, Exscientia will be able to go to regulators with a drug that consistently works well in a very defined patient population.

"This is where AI is going to lead as well. It's not just about using AI to make drug discovery better, but about how we can create better drugs overall," Hopkins says.

In January 2023, Insilico Medicine announced an encouraging topline readout of its phase 1 safety and pharmacokinetics trial of INS018_055, designed by AI for idiopathic pulmonary fibrosis, a progressive disease that causes scarring of the lungs. Their proprietary AI platforms identified a new target (which Zhavoronkov would identify only as 'target X') and a small molecule inhibitor, which was granted breakthrough status by the Food and Drug Administration (FDA) in February.

"It's the first time anyone in our industry has developed a novel target of a molecule, and completed phase one trials, all the way with AI," Zhavoronkov says. He expects phase two readouts in the first half of 2023. It is part of Insilico's growing pipeline targeting diseases associated with aging. What makes Insilico's work more impressive, according to Zhavoronkov, is that the company only began development on INS018_055 in February 2021.

"We have 31 therapeutic programs. In 2020, we had zero," says Zhavoronkov.

AI for analysis

Recursion, a biopharma startup based in Salt Lake City, Utah, uses AI not to design molecules but to analyze data from millions of experiments and billions of microscopy images that their lab is gathering with the help of robots.

"Just like Google has all these cars driving around taking pictures that they turn into really useful maps for all of us, we've done the same thing with biology," says Chris Gibson, Recursion's co-founder and CEO.

Recursion is also working to develop a therapeutic agent for ovarian cancer that targets a gene that their AI systems indicated was part of the same pathway as CDK12, an existing target that has proved challenging to inhibit directly. In preclinical studies that target the CDK12-associated protein, 40% of mice showed a complete response. When the compound was paired with a PARP inhibitor, tumors were eliminated in four out of five mice. The company also has three other compounds in clinical trials for oncology and rare diseases: familial adenomatous polyposis, cerebral cavernous malformation, and neurofibromatosis type 2.

"Biology and chemistry are so broad and complex. Your goal isn't to find everything. Your goal is to find something really good and advance it," Gibson says.

Relay Therapeutics has developed an oral, small molecule inhibitor of FGFR2, a receptor tyrosine kinase that is overactive in certain cancers, such as intrahepatic cholangiocarcinoma. Existing FGFR inhibitors are not very selective, but the company is testing RLY-4008, which is only active against FGFR2. At the end of 2022, BenevolentAI completed a phase 2a trial for BEN2293, a topical

ointment for the treatment of atopic dermatitis (eczema). The treatment was found to be safe but did not meet its secondary endpoint of reducing itch and inflammation, according to a company press release in April 2023.

BenevolentAI has also filed a clinical trial application with the UK Medicines and Healthcare Products Regulatory Agency (MHRA) for BEN-8744, a small molecule phosphodiesterase 10 (PDE10) inhibitor designed to treat ulcerative colitis. If approved, Williams-Jones says BenevolentAI plans on beginning a phase 1 trial in the first half of 2023. But for BenevolentAI, as for everyone else, he points out this is still early days.

"Biology is hard, and we don't know very much in real terms," says Williams-Jones. Every time scientists think that they have made a big step forward in simplifying the drug development process, he says, they stumble across two or three other issues that they did not expect.

The protein folding problem

Much of AI-driven drug discovery builds on protein folding. By the latter half of the twentieth century, biochemists had decoded some of the basics of protein structure tenets that now fill biology textbooks. A string of amino acids, proteins fold into complex, three-dimensional structures based on the atomic interactions between the backbone and amino acid side chains. This structure determines the protein's function. As crystallography and electron microscopy began to crack open the atomic-level structures of proteins, biochemists began to wonder whether it might be possible to predict the final structure of a protein complex using only its amino acid sequence. The discovery of α-helices and β-sheets in the 1960s made the promise seem almost tractable.

Then reality began to sink in. Twenty simple amino acid building blocks could give rise to a dizzying array of proteins – greater than the number of stars in the universe, says David Baker, Head of the Institute for Protein Design at the University of Washington. Methods such as multiple sequence alignment (MSA) enabled structural bioinformatics experts to compare the amino acid sequences of numerous protein homologues to determine domains, disordered regions, and other elements of local secondary structure. But even the most advanced MSA methods could not reveal allosteric interactions, or how different α-helix regions were arranged next to each other.

AI and machine learning took a completely different approach. "Machine learning is based on the results you attain rather than a statistical model that describes the population," Deane says. "It's about finding predictive patterns in the data."

Instead of applying the laws of physics to every single atom or bond, what if scientists began to look for similarities between proteins? If they could assemble a reasonably broad base of protein structures (gathered the old-fashioned way, through painstaking crystallography, X-ray diffraction, and electron microscopy techniques), then perhaps scientists could try to figure out the similarities between proteins and use that to predict a protein's structure.

"With deep learning, you don't really try and simulate the actual folding process. You're not trying to find the lowest energy state. It's more about pattern recognition," Baker says.

The intellectual leap to this way of thinking was profoundly important, says Alan Lipkus, senior data analyst at Chemical Abstracts Service in Columbus, Ohio.

Weird molecules

By the early 2010s, computer scientists and computational chemists had developed the prototypes of groundbreaking AI systems such as RoseTTAFold and DeepMind's AlphaFold. Most modern machine learning algorithms devoted to predicting protein structure contain four different modules: an input module that contains the amino acid sequence and structures from homologous proteins; a sophisticated neural network that uses pattern recognition algorithms to transform the amino acid sequence into spatial information of the protein; an output module that converts the spatial information into a preliminary three-dimensional structure; and a refinement process that enables fine-tuning. Using these algorithms, AlphaFold2 can predict single protein domain structures down to 2.1 Å, essentially solving the protein structure problem. It is a staggering accomplishment, Baker says, but he wants to move beyond it.

"By just predicting protein structure, you're stuck with whatever exists in nature. You can't make anything new. But now we can make all these brand-new proteins for cancer therapeutics and clinical trials. You can make all kinds of different things with protein design," Baker says.

Beyond the basic science accomplishment, these advances have also given a huge leg up to pharma. Determining a protein's structure was a major hurdle in designing the right molecule to alter its function. Determining the structure of a small molecule was simple compared to a protein. Even biologics designed by AI were a possibility, antibodies just being one specific type of protein. This progress did not remove the need for experiments and tinkering – no computer algorithm is yet that good – but it narrowed down the number of possibilities to help scientists prioritize molecules that were far more likely to have the desired effect without causing undue toxicity.

The molecules these AI systems helped to design, however, looked very different from compounds designed by medicinal chemists. When InSilico's Zhavoronkov began pitching his AI therapeutic design service to pharma companies, he included examples of several molecules his system had built. Their novelty immediately grabbed the attention of potential pharma partners, some of whom helped provide series A and B funding rounds.

"They said to me: Alex, these molecules look weird. Tell us how you did it," Zhavaoronkov says. "We did something in chemistry that humans could not do."

And it is this weirdness that just might be AI's biggest strength in pharmacology. Although the total number of possible chemicals in the universe – what some scientists refer to as chemical space – is vast, humans have only explored tiny slivers of this space. Synthetic chemists develop expertise working with certain types of compound or performing specific reactions, says Lipkus, leading to a few small areas of chemical space that are well mapped out. Most of chemical space remains terra incognita.

Many clinical trials test tweaks of existing drugs, which may give a slightly improved safety or efficacy. However, a much bigger prize is a first-in-class drug against an entirely new target, which AI-designed drugs are well-positioned for.

Lipkus and his colleague Todd Wills (now a senior vice president at Cass Information Systems) analyzed the novelty and creativity of pharmaceutical molecules using the chemical abstract service database of thousands of molecules, which "is probably the best representation of they known chemical universe", Lipkus says. They compared the uniqueness of a molecule's scaffold and shape, which they defined as the atom-to-atom connectivity that prunes back all but the most basic information about a compound's structure. 'Me too' drugs, they pointed out, tend to consist of small alterations to a drug's chemical side chains rather than large-scale shifts in molecular structure. A growing number of pharmaceutical compounds, they pointed out in a 2019 paper in the *Journal of Organic Chemistry*, are showing signs of creativity, with more unique

structures and scaffolds. AI, Lipkus says, will only accelerate this trend.

"It's one more piece of evidence that there's value in looking for novel structures," Lipkus says. "Talking to people in the drug industry, they want to break away from these scaffolds that have been used so heavily."

AI tools also enable drug developers to explore the chemical world much more quickly.

"It allows us to explore a much broader slot or chemical space than we'd be able to using experimental methods on their own," says Don Bergstrom, president of research and development at Relay Therapeutics.

Neglected diseases

AI-designed drugs are not just being developed for potential blockbuster status. In Geneva, Switzerland, the Drugs for Neglected Diseases Institute (DNDi) is using machine learning to create better drugs for conditions that predominantly affect the world's poor, such as Chagas disease and dengue fever. Charles Mowbray, discovery director at DNDi, says the institute is also turning to AI strategies to guide its drug repurposing pipeline as part of its global efforts to develop therapies for neglected diseases. For such diseases, speed is critical; AI can help scientists generate hypotheses and test them more quickly.

"These tools don't replace a scientist, they complement them," Mowbray says. "[AI] enables them to have all the information at their fingertips, to ask the good questions, to refine their queries, and to iterate until they can figure out what they're really after." This synergy is true for machine learning across drug development, he adds.

Even as the impacts of AI in drug design are beginning to emerge in clinical trials, these strategies are joining other AI tools in clinical trial design, manufacturing, and more. There is no doubt that machine learning is profoundly reshaping the pharmaceutical industry, Lipkus says. As for how the effects of AI-developed drugs will play out, he is more circumspect, saying that is still up in the air.

"Nothing guarantees anything. Drug discovery is really difficult. I don't know if people expect AI to just pop out the design of a molecule that's your next blockbuster, says Lipkus. "It's all kind of a crapshoot."

Carrie Arnold
Science writer, Richmond, VA, USA.

Published online: 1 June 2023