Global Insights

Smart Thinking on AI in Healthcare Part 4: Biopharma

AI is likely to disrupt Biopharma, moving it away from "drug discovery" and toward "drug design"

23 Jun 2023 12:47:19 ET

Executive Summary

Adam Spielman
+44-20-7986-4211
adam.spielman@citi.com

Wenyan Fei
wenyan.fei@citi.com

We believe AI is especially important for biopharma, because AI can understand biology in a way humans cannot. The reason that ChatGPT is so impressive is that the underlying Transformer technology can "understand" ordinary languages like English. Transformers can also "understand" large bio-molecules like proteins. Traditionally drug development has involved informed trial and error to *discover* molecules, followed by long, expensive clinical trials, in which more than 90% of drugs fail. With AI drug development can move closer to drug *design*, speeding up the process and creating better-targeted drugs. Clinical trials can move from *exploring* the properties of a molecule that has been discovered to *confirming* that the design works as intended. This is likely to shorten their duration, and make the likelihood of success much greater. Biopharma moves slowly, but over time we believe it will be disrupted more by AI than almost any other industry.

"We are reimagining the entire drug discovery process from first principles with an AI-first approach"

– Demis Hassabis, founder and CEO of Deep Mind, and CEO of Isomorphic, an Alphabet subsidiary, which aims to bring AI to healthcare.

"The use of artificial intelligence and data science already support our teams' efforts in areas such as accelerating drug discovery, enhanced clinical trial design, and improving manufacturing and supply of medicines and vaccines. We have just scratched the surface as to how we embrace these disruptive technologies."

– Paul Hudson, CEO of Sanofi

"AI doesn't represent an incremental change; it's going to fundamentally change how drug discovery works"

– Neal Batra, Deloitte Consulting Principal and Future of Health leader.

"What I get most excited about is the potential for AI to fundamentally change the timescales, as well as the cost structures, for development in the therapeutic space."

– Vineeta Agarwala, Andreessen Horowitz General Partner focussing on therapeutics, diagnostics and digital health at the bio+health fund

Contents

The first section of this report explains why we think AI will have such a profound impact on biopharma.

The second, starting on page 24, explains what is already happening in some detail.

Note on terminology: Biopharma, biotech and pharma

Biopharma is the term used to encompass both the *pharma* and *biotech* industries.

- **Biotech** companies focus on developing drugs and taking them through clinical trials. These companies usually have little or no revenue, but substantial expenses, and are therefore heavily loss-making. Many of these companies disappear without ever producing viable medicines.

- **Pharma** companies sell approved drugs, and therefore have revenue and are usually profitable. The issue for them is there is a patent on every drug, and when it expires, the profit contribution usually collapses as generics enter the market. The pharma companies therefore also develop new drugs, and frequently buy promising biotech companies too.

In practice there are about 20 large pharma companies globally. By contrast about 1,800 companies have sponsored FDA-registered clinical trials for new drugs between 2010 and 2020 – of which about 750 sponsored only one candidate drug.[1]

There is, unsurprisingly, a continuum between the two terms, and sometimes it is a bit arbitrary if a company is called a “pharma company” or a “biotech”. We therefore use the term *biopharma* to capture the entire range of companies.

[1] Source: Biomedtracker

Smart thinking on AI in health: the impact on Biopharma

RELATED:

Smart Thinking on AI in Healthcare: Part 1 – Overview
Smart Thinking on AI in Healthcare: Part 2 – Medical Devices
Smart Thinking on AI in Healthcare: Part 3 – Health Administration

In our view, AI will change biopharma as profoundly as the promise of electric and autonomous vehicles is changing the auto industry

This report explains why we think biopharma will be affected by AI more than any other aspect of the healthcare industry, and in fact more than almost any other industry. It is the fourth report in our series *Smart Thinking on AI in Healthcare.*

We believe both that (1) AI is especially important for biopharma, because AI can understand biology in a way humans cannot, and that (2) biopharma is in a particularly strong position to adopt AI, because it is simultaneously highly profitable and extremely inefficient (given the iterative trial and error approach towards drug discovery).

Figure 1. Reports published and planned in the Smart Thinking on AI in Healthcare series

1	**Overview**	AI will progressively transform the health system
2	**Medical Devices**	AI devices are already becoming more common, especially in radiology
3	**Health Administration**	AI will help automate a lot of admin, taking over mundane tasks and saving time and cost
4	**BioPharma**	BioPharma will be deeply affected as transformers analyze large molecules
5	**Role of Doctors**	Gradually AI will automate diagnosis and prescription, letting doctors focus on higher tasks

Source: Citi Global Insights

1. AI is especially important for biopharma

The reason AI is likely to have such a profound impact in biopharma is that the same sort of transformer technology that can (sort of) understand natural languages like English can also be applied to long molecules like proteins and RNA. Humans understand and can manipulate English better than AI can; by contrast the way molecules work is not intuitive to humans. As a result it's fair to say that in many ways AI understands the language of large bio-molecules better than humans do.

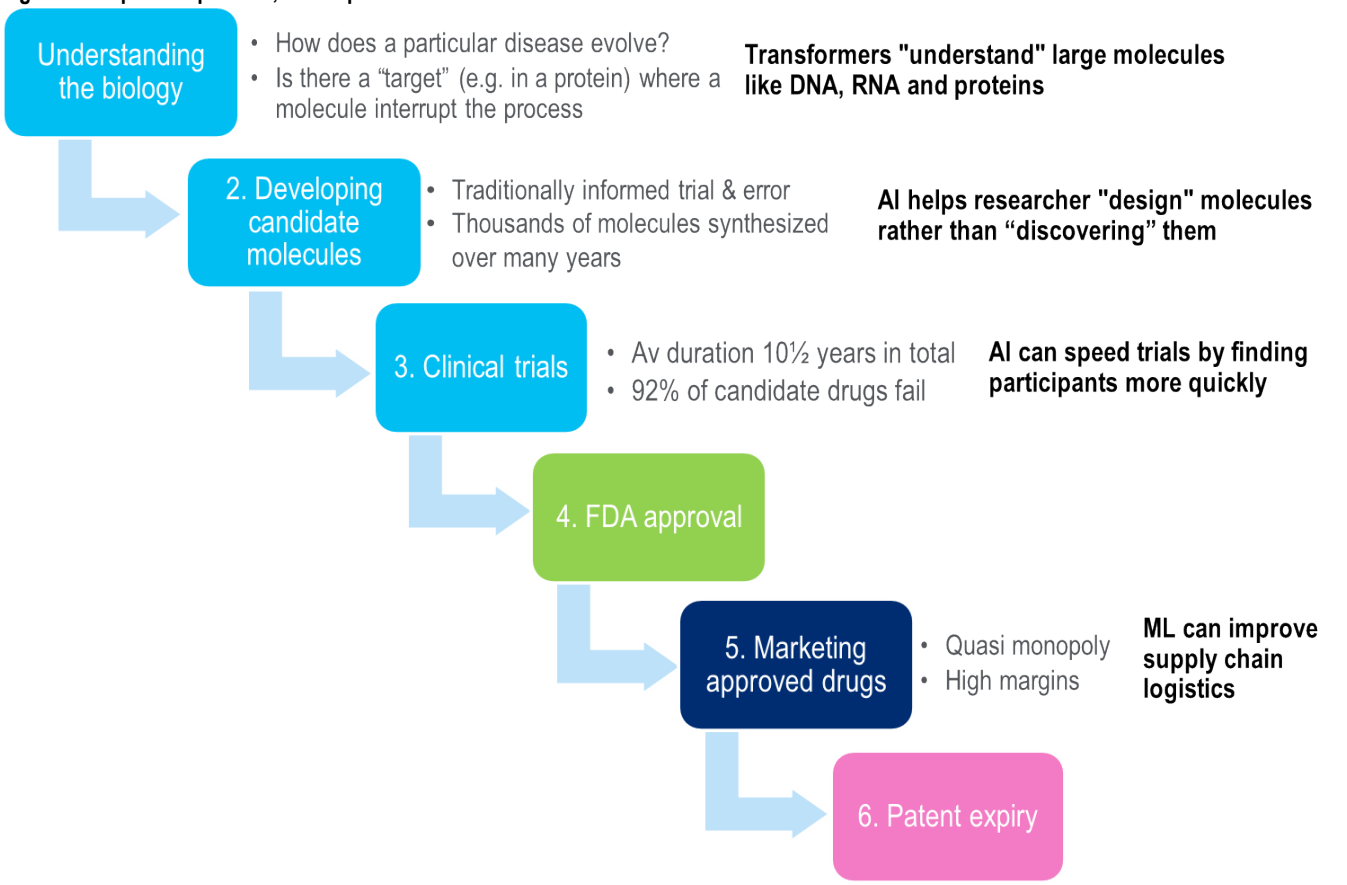
One sign of AI's power to change biopharma is that several of the individuals who have been at the cutting edge of AI are now focusing on using AI to accelerate the development of drugs – for example Demis Hassabis (the founder and CEO of DeepMind) and Daphne Koller (a former professor computer science at Stanford).

Another indication is that of all the academic papers on AI, the ones that have been most cited in the past couple of years have actually focused on proteins.

AI's biggest impact is likely to be in developing molecules but it can help with almost every aspect of biopharma

In very simple terms biopharma works in six stages, as set out in Figure 2. AI's biggest impact is likely to be in understanding the biology (stage 1) and developing candidate molecules (stage 2), but it can also help make clinical trials more efficient and improve supply chain logistics.

Figure 2. Biopharma process, and impact of AI



Source: Citi Global Insights, Biomedtracker and Pharmapremia

2. Biopharma is also in a very strong position to adopt AI

One reason why the impact of AI on biopharma is important is that it is the dominant subsector within healthcare in financial terms -- biopharma accounts for roughly half of the healthcare industry's profit and investment

Biopharma operates a business model quite unlike any other industry's, and many of the peculiarities seem tailor-made to adopting AI:

1. **Lots of cash to invest in AI.** When drugs are approved they operate as quasi-monopolies, with very high margins. This means both that many pharma companies have strong cash flows, and that there are huge levels of R&D investment in the biopharma sector.

2. **There is plenty of scope for AI to bring improvements because the industry involves extraordinary inefficiencies.** Developing new drugs often takes 15 years and costs billions of dollars. Typically thousands of molecules are tested before one is taken to clinical trials – but even then 92% of the candidate drugs fail in clinical trials. For those drugs that succeed it takes an average of more than a decade from the start of clinical trials to approval, as shown in Figure 2.

3. **And time is of the essence.** Biotech companies have to file patents on their drugs before they start clinical trials, and the patents last for only 20 years, at least in theory. Even though they are often extended in practice, pharma companies have only a limited number of years to generate returns from the drugs. Anything that can bring drugs to market faster is therefore likely to have a major economic impact – and there is plenty of evidence that AI should be able to do that.

Moving from "drug discovery" toward "drug design"

The reason it costs so much, and takes so long to develop new drugs, is that biology is incredibly difficult. The industry talks about "drug discovery", and it does involve a lot of iterative trial and error, with researchers slowly closing in on candidate drugs. In some ways it is like trying to find a particular rock, when you are searching across an entire continent, and only have a few, not very accurate maps.

AI has the power to change that, because it promises to radically improve our understanding of biology, and as a result developing new drugs will be much closer to "design" than discovery. Researchers now have much better, multi-dimensional (albeit still imperfect) maps to guide their search.

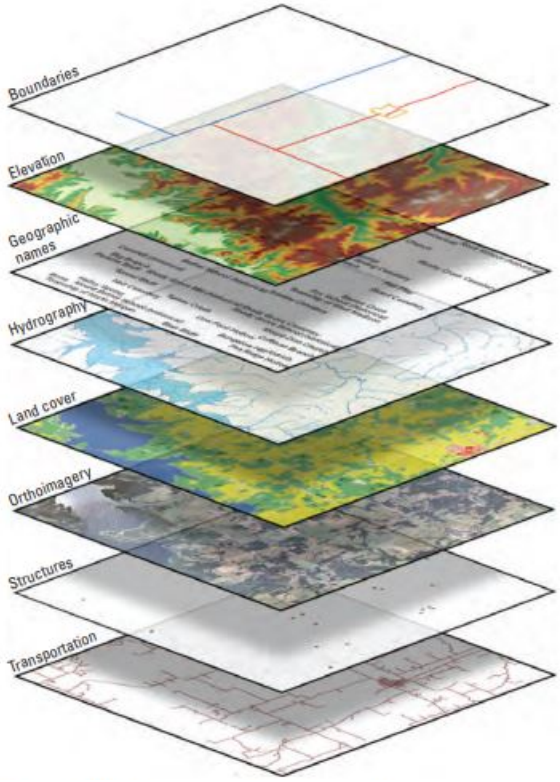
"If you look at the history of drug development, particularly research, it's largely been a lot of trial and error," says Greg Myers, chief digital and technology officer at Bristol Myers Squibb.[2] " What we're really hoping is that software, machine learning and deep learning in particular are ways to try to operationalize the serendipity of what we're used to seeing."

[2] https://bio-eats-world.simplecast.com/episodes/adapting-biopharma-to-ai-Uh74PslCFFRFG

Figure 3. Finding a rock, when you are searching across a continent with not-very accurate maps

Source: USGS

Figure 4. AI has the potential to create more accurate (but still less than perfect) biological mapping



Source: USGS

Cancer drug designed with AI

Relay's anti-cancer drug (RLY4008) is another molecule that was designed by AI and is now being tested in the clinic.
See page 30.

One example of a drug that has been designed by AI is EXS21546 (made by **Exscientia**). This is now in Phase I/II trials, focusing on kidney cancer (RCC) and lung cancer (NSCLC). The drug is a type of A_{2A} receptor antagonist. Researchers believe A_{2A} receptor antagonists are the best treatment for people with tumors that have high concentrations of adenosine[3]. Exscientia's analysis predicts about 20%-50% of RCC and NSCLC patients fall into this group – so A_{2A} receptor antagonists are potentially important. They work by restoring anti-cancer immune activity provided by T-cells that the excessive concentrations of adenosine has inhibited. Until now, however, it has proved hard to make acceptable A_{2A} receptor antagonists because they tend to cause severe side effects, notably in the brain.

The use of AI-design has certainly made the development faster. It also appears to have made it more accurate:

- **Faster:** Exscientia identified this particular drug ('546) within nine months, by using successive AI-driven design cycles, and after testing only 163 compounds. This was less than 20% of the time taken during the average (traditional) discovery process (54 months) -- and required synthesis of less than 10% of the averages number of molecules (2,500).

[3] Adenosine in one of the four building blocks of RNA.

- **More accurate:** In vivo, '546 demonstrated both effectiveness (it reversed adenosine-related immune suppression) and selectivity (in other words it didn't affect brain cells.) In a Phase 1 study – focused only on healthy volunteers – the molecule confirmed the company's design, with good selectivity with no adverse events in the central nervous system.

The next clinical trial is now recruiting.

Transactions for drugs developed with AI

Certain drugs developed with AI have been sold for substantial sums. A drug designed with AI by **Nimbus Therapeutics**[4], for example, was sold this year to **Takeda** for $4 billion cash, plus up to a further $2 billion in additional payments. The drug (NDI034858) is targeted at various autoimmune diseases, including psoriasis. It has done well in Phase I and II trials but still have to be tested in Phase III.

In fact this is the second time Nimbus has sold an AI-designed drug for a significant sum. In early 2016 it sold Firsocostat, a drug for NASH, which is a type of liver disease, for $400 million upfront, and upto $800 million in milestone payments to **Gilead**.

Ultimately, however, the drug failed in Phase III – a useful reminder that no matter whether a drug is designed by AI, and no matter how promising the results are in Phase I and II, nothing is certain until Phase III trials are complete.

Improving the chances in clinical trials

Clinical trials are necessary to show drugs are both effective and safe. However they are the most expensive and time consuming part of developing new drugs[5] – and more than 90% of drugs fail during the trials. But AI has the power both to speed them up and to increase the likelihood of success.

First, AI-first drug development alters the nature of Phase I and II trials, making them more about *confirming* that the drug works as it is designed, and less about *exploring* its properties. Secondly, AI can help clinical trials in very practical ways– for example finding more participants more quickly, including from historically under-represented populations. Thirdly, AI can be used to find patients who are more likely to respond well to the drug.

See page 38 for more details on Vida Lung

Vida Lung is an example of a company applying AI-driven image recognition analysis to X-rays and CT-scans of lungs. The company aims to recruit the right participants quickly and track changes in lung structure accurately over time.

"Clinical trials have traditionally been slow and inefficient, which is why they're so expensive," Susan Wood, the CEO, told us. "If we can make them more efficient – by finding the right patients and sites quickly and measuring treatment response more precisely – that creates significant value."

---

[4] Nimbus is a private company based in Boston. It uses AI to design small molecules to bind with well-validated but difficult-to-drug targets. The aim is to create therapies for widespread diseases that have – until now – proven difficult for drug makers to tackle.

[5] Various estimates suggest that clinical trials account for 50%-60% of the cost of developing new drugs. See https://www.frontiersin.org/articles/10.3389/fmed.2021.760762/full#B1

This third point -- about using AI to find participants who are likely to respond well to a certain drug -- is illustrated by the **EXS21546** example we've just discussed. For clinical trials of A_{2A} receptor antagonists it's important to find patients with tumors that have high concentrations of adenosine, because they're the people the drug is designed to work for. Until now, however there hasn't been any practical way of identifying them.

Exscientia, however, has created an "adenosine burden score" to help find the relevant patients. The score works by measuring gene expression levels that corresponded to high adenosine levels - and the score was originally found by using complex primary model systems and single cell technologies. Exscientia's approach should allow it to find the right participants and therefore increase the likelihood of a successful outcome.

AI can also help in the supply chain

AI can also help improve manufacturing and logistics, even though this is the least important part of the pharma industry (in economic terms.)

Sanofi for example says it is using AI to improve batch yields and to predict areas with low stock levels.

“The use of artificial intelligence and data science already support our teams’ efforts in areas such as accelerating drug discovery, enhanced clinical trial design, and improving manufacturing and supply of medicines and vaccines,“ said Sanofi’s CEO, Paul Hudson. ”We have just scratched the surface as to how we embrace these disruptive technologies to achieve our ambition of transforming the practice of medicine.”

Reasons for a measured view

As we write throughout this report, we are confident AI can change the biopharma industry profoundly. However, it's important to realize that even with AI the process of creating drugs will remain slow and expensive.

Long-lead times

AI's impact is likely to be gradual, because the industry operates on very long lead times.

Clinical trials are the only way of truly demonstrating a potential drug is safe and effective, and even with AI, they will often take many, many years, as we discuss in detail, starting on page 37.

It's true that the trials for the Covid vaccines were conducted in about a year, but it would be wrong to draw general conclusions from this: (1) Covid's progression is measured in weeks; and (2) it was extremely common. By contrast many conditions evolve over decades and clinical trials need to reflect that. Furthermore, for many diseases, recruiting participants is a slow affair.

And of course developing a drug with AI does not guarantee that it will succeed in clinical trials. As we have already mentioned, the drug Firsocostat, which Nimbus developed with AI and which Gilead bought, ultimately failed in Phase III.

Not a silver bullet

Another reason for caution is that previous approaches that promised to make drug development more effective haven't made much difference in practice. Many people argued that combinatorial chemistry would accelerate drug development, for example, but the technique largely failed to live up to its advocates' hopes.[6]

Vineeta Agarwala is general partner at A16Z's bio + health fund, a significant investor in AI-driven biopharma: "I think what people get wrong is to assume that AI is a catch-all silver bullet for problems that are hard. Those problems will still be hard," she said in an interview.[7] "There are vexing problems of target biology, of interconnected pathways that we don't understand yet, of preclinical models not being fully predictive of human biology. AI will create insights that make those problems a little bit more tractable, more scalable, more engineerable. But many of the problems won't be solved by AI alone."

[6] https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3564354/

[7] https://www.mckinsey.com/industries/life-sciences/our-insights/vineeta-agarwala-on-the-promise-and-limits-of-ai-in-drug-discovery

Implications for the industry

A whole raft of new companies have already grown up around AI in biopharma. Some are helping the megacap pharma companies, with a focus on particular aspects of biology, or drug discovery. Other companies are trying to use AI to develop drugs themselves, from soup to nuts.

Pharma

The existing megacaps are, of course aware of the trends, and are incorporating more and more AI into their business models. Their financial resources mean they will have the capacity to adapt, but it will likely be a big cultural change.

Biotech

We believe adapting will be even more important – and probably harder – for the long tail of biotech companies, since they are much smaller than the pharma companies, and more resource constrained. In theory, however, AI could be of more help them more as it has the power to make developing drugs faster and cheaper.

If this sector can adapt successfully, it should help accelerate the number of drugs that come to the market. Even without the impact of AI, more than 50% of new-drug approvals in the U.S. – and more than 60% of blockbuster drug[8] approvals – come from biotech companies with sales below $500 million.[9]

AI likely to have profound financial implications

AI is likely to improve the industry's cashflow:

- **New drugs can generate revenues in the billions.** If AI allows new drugs to be found – to drug the previously undruggable – then the upside could potentially be worth billions of dollars.
- **Bringing drugs to market faster can improve cash flows very substantially.** AI can produce drugs more quickly, which means the drug is likely to produce revenue for a greater percentage of the patent period. Accelerating approval by 12 months can increase the NPV of a drug by $100 million.[10]
- **Better drugs can generate higher revenues.** AI-designed drugs have the potential for greater revenue, even if they are treating diseases where therapies already exist, because they can be targeted more tightly and designed with fewer side effects.
- **AI can reduce the cost of development substantially**, especially in the discovery phase. Exscientia says AI-driven drug development can reduce the costs in the preclinical phases by about 70%, relative to traditional techniques.

[8] A blockbuster drug is a drug with annual sales of more than $1 billion

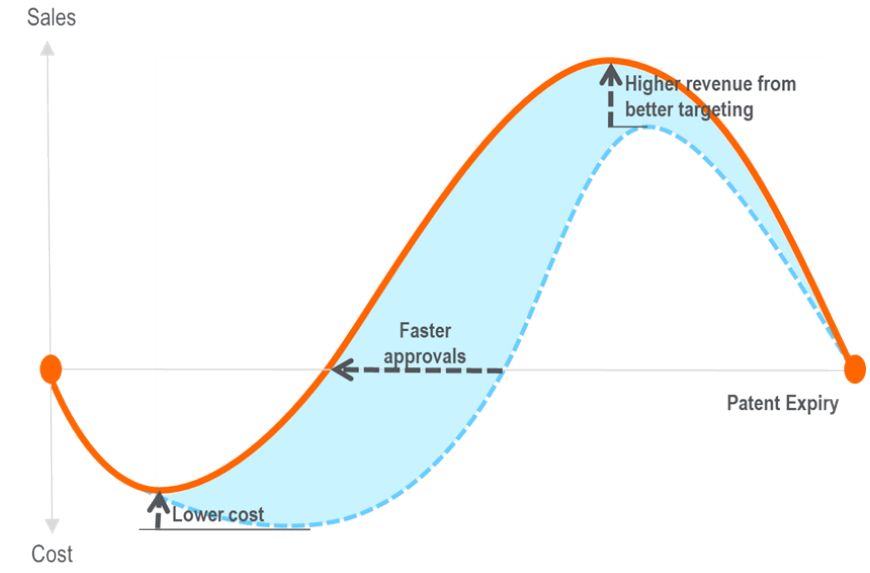
[9] https://www.iqvia.com/insights/the-iqvia-institute/reports/emerging-biopharma-contribution-to-innovation

[10] Source: AbCellera

Computer science

The leading AI-first biopharma companies already employ as many computer scientists as life scientists. We think that in future this will become more and more common, and most biopharma companies will be focused as much on computer science and as on life science.

Figure 5. Cash Flow from AI design drugs vs traditional design



Source: Exscientia, Citi Global Insights

Supplying AI as a service vs. developing a pipeline

For the AI-driven start-ups, there is a choice – whether to focus on only one area of the drug development process, suppling AI-as-a-service to the large, established biopharma companies, or whether to try to become a full-stack biopharma company, developing drugs and taking them through clinical trials.

One of the largest AI-first companies (**AbCellera Biologics**) is very much in the AI-as-a-service model. AbCellera focuses on using AI to design antibodies and aims to be the "partner of choice" in the field. By the end of March, the company was working with 41 discovery partners under 177 contracts, and it had nine molecules in clinical trials.

By contrast a company like **Exscientia** believes the greatest competitive advantage will come from a full-stack approach. The company defines its mission as "encoding and automating every stage of drug design and development."

"The big question is who is going to be able to use AI really effectively first," Ben Taylor, Exscientia's CFO and chief strategy officer, told us. "It's about integrating systems --every stage of the system has to be able to feed into all the others. If you get that right, and create a closed loop, then it's massively scalable. There will be a huge difference between those who can create a genuine learning system and those who can't."

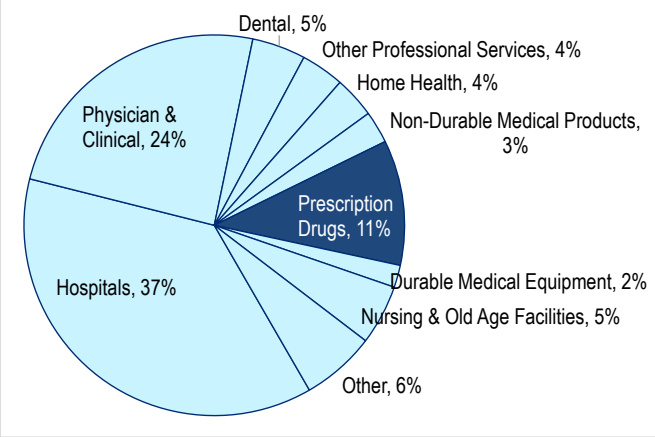
Nate Benaich, General Partner at Air Street Capital, agrees. “AI-first drug discovery companies are set up to directly monetize the outputs of their platforms - novel medicines. By contrast, selling drug discovery as a service would leave the vast potential upside of blockbuster drugs on the table."

Biopharma is the largest subsector within healthcare in economic terms

One reason why the impact of AI on biopharma is important is that it is the dominant subsector of healthcare in financial terms. Biopharma accounts for very roughly half of the industry's profit and investment, even though prescription drugs account for only 11% of spending on healthcare in the U.S. (Figure 6.)

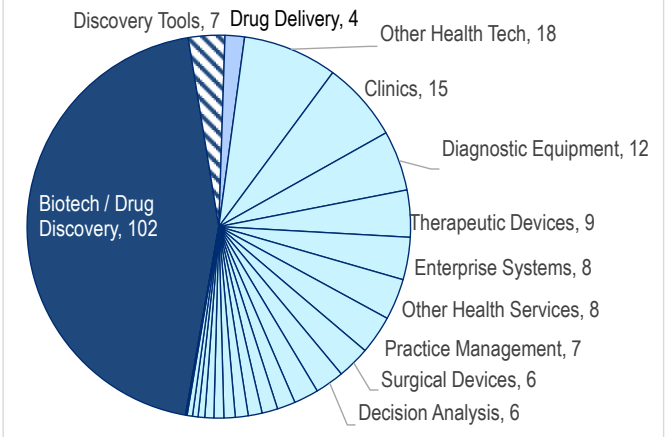
Figure 7 shows that over the last two years, $102 billion was invested in biotech and drug startups, and a further $7 billion was invested in tools for drug discovery, together making up 48% of total investment by VCs in healthcare.

Figure 6. U.S. Healthcare Spend by Category, 2021



Source: CMS

Figure 7. Primary Capital Raised in Healthcare Startups, 2021 and 2022 combined ($ in Blns)



Source: PitchBook

Figure 8 by contrast shows the split of market cap within the publicly listed health companies. Globally, 41% of health market cap is accounted for the pharma companies, and 15% by the biotech sector. Medical devices and healthcare providers are much smaller in terms of market cap.

Figure 8. Healthcare sector market cap with the public equity market, split by subsector

	Global	US	Europe
Pharma	41%	29%	77%
BioTech	15%	18%	3%
Life Science Tools	11%	13%	6%
Health Equipment	18%	19%	13%
Providers	16%	21%	2%
Total Healthcare	100%	100%	100%

Data as of 16 April 2023
Source: FactSet

Which companies are using AI in biopharma?

Please contact the Citi Global Data Insights team to request screening of companies exposed to the themes in this report

Our colleagues in Citi Global Data Insights are able to respond to requests to generate screens of companies exposed to the themes in this report, based on a quantitative analysis both of patents filed and of newsflow.

These screens can cover quoted and unquoted companies. They can be ranked in many ways. For example it is possible to rank by the number of patents obtained related to AI in biopharma, or the quality of those patents, or the percentage of a company's patents that fall in this area.

Please do reach out to Helen Krause (helen.krause@citi.com), for more information.

AI is especially important for biopharma

AI comes in many forms, as we discussed in AI Time - 10 Ways AI is Getting Real. So far, the main types of AI that have impacted healthcare have revolved around image recognition – which is why there has been such a big impact on radiology – and machine learning, which is vital for health administration and automated diagnosis.

Transformers are among the most important developments of the past few years

However we believe it will be "Transformers" that revolutionize biopharma. Transformers were introduced in 2017, with a paper called "Attention Is All You Need", written by Google Brain.[11]

Transformers take any sequential information – for example the words in a text, or the amino acids in a protein – and analyze them by taking into account the relationships between the elements that come before and after, weighing them by their importance.

"Unlike previous [techniques] that processed words individually in order, transformers understand the context and relationship between each word and all the words around it," explains Jeffrey Zhu, a project manager at Microsoft.

The same principles can be applied to large bio-molecules."Proteins are a lot like human language -- they can be naturally represented in strings and letters," points out Greg Myers, the chief digital officer at Bristol Myers Squibb. "Protein motifs and domains act a lot like the functional building blocks of a language . . They're akin to words, phrases and sentences."

It's now reasonable to say that what are called Large Language Models, or LLMs, are getting better at *understanding* languages like English, Chinese and Spanish. AI can now extract the main points from an unstructured piece of text, and also write in an impressively human-sounding way, as ChatGPT and GPT-4 show.

Humans are, of course, fluent in ordinary languages, and so we tend to leap on the weaknesses in programs like ChatGPT. By contrast, humans do not naturally understand what the coding within a strand of DNA or a protein means, and therefore any understanding – even if it is incomplete -- is a big step forward.

Moravec's Paradox and drug development

Moravec's Paradox states that it is often easy to get computers and robots to do things that humans find hard, like long division and playing chess, and much harder to get them to do things that humans find easy, like walking across a crowded street. Understanding biology and developing drugs are clearly in the category of things that humans find hard. We think they are also moving into the category of things where computers hold a significant advantage.

[11] https://arxiv.org/abs/1706.03762

Protein folding

One celebrated example of the use of AI in biology is the group of tools that can now predict the 3D shape of protein molecules, for example AlphaFold, developed by DeepMind, and RoseTTAFold, developed by the University of Washington.

Previously the only way of knowing the shape of a protein was using X-ray crystallography, a technique that goes back to early in the 20th century. And knowing the shape of a protein is often critical, because most drugs involve getting a small molecule to bind onto unwanted proteins, to interrupt the process they are causing that leads to disease.

One example of the way this technology is helping is in developing a vaccine for malaria.

Last May year researchers at Oxford announced they had used AlphaFold to establish the structure of a protein involved in the transmission of malaria, allowing them to say where antibodies that could block transmission are likely to bind.[12]

"Previously, we'd been using a technique called protein crystallography to work out what this molecule looks like, but because it's quite dynamic and moves around, we just couldn't get to grips with it," Professor Matthew Higgins, the lead researcher, said.[13] "When we took the AlphaFold models and combined them with this experimental evidence, suddenly it all made sense. This insight will now be used to design improved vaccines which induce the most potent transmission-blocking antibodies."

Several leading figures within AI have moved into biology

One of the signs that AI really does have the power to revolutionize biopharma is that several of the individuals who have been at the cutting edge of AI, for example **Demis Hassabis, Daphne Kollar** and **Jakob Uszkoreit**, are focusing on applying AI to drug research:

- **Demis Hassabis** is co-founder and CEO of DeepMind, the company that developed both AlphaGo and AlphaFold. In 2021 he started Isomorphic Labs, with the aim of "reimagin[ing] drug discovery from first principles with an AI-first approach", and he remains CEO.
- **Daphne Koller** was professor of computer science at Stanford and a MacArthur Fellow[14]. She cofounded Coursera in 2012, but her new company is **insitro**, which uses ML and generating biological data "at scale" to help develop drugs. Among other projects, insitro is working with **Gilead** to create experimental models for NASH (a type of liver disease) and with **Bristol Myers Squib** on neurodegenerative disorders.

---

[12] bioRxiv May 2022: Structure of a malaria vaccine candidate and its recognition by transmission blocking antibodies

[13] https://www.theguardian.com/technology/2022/jul/28/deepmind-uncovers-structure-of-200m-proteins-in-scientific-leap-forward

[14] The MacArthur Fellows program is popularly called "the Genius Grant". It is granted to about 20-30 U.S. residents a year for "extraordinary originality" in their field.

- **Jakob Uszkoreit** has spent much of his life on machine translation. He is a co-author of the foundational paper on transformers (Attention is All You Need, 2017) and formerly head of the Google Brain in Berlin. His new company is **Inceptive Nucleics**, which is focused on designed new types of RNA through a combination of highly scalable experiments and deep learning.

The most read papers within AI have focused on proteins

Our view that AI is particularly well suited to drug discovery is also supported by the citation statistics: Despite the excitement around ChatGPT, Figure 9 shows that of all the papers on AI, the ones that have been most cited in the past couple of years have actually focused on proteins.

Figure 9. Top 3 academic papers mentioning AI, rated by citations, 2021 and 2022

	2022
1	AlphaFold Protein Structure Database: massively expanding the structural coverage of protein-sequence space with high-accuracy models DeepMind, 1372 citations Using AlphaFold to augment protein structure database coverage
2	ColabFold: making protein folding accessible to all multiple institutions, 1162 citations An open-source and efficient protein folding model
3	Hierarchical Text-Conditional Image Generation with CLIP Latents OpenAI, 718 citations DALL·E 2, complex prompted image generation
	2021
1	Highly accurate protein structure prediction with AlphaFold DeepMind, 8965 citations AlphaFold, a breakthrough in protein structure prediction
2	Swin Transformer: Hierarchical Vision Transformer using Shifted Windows Microsoft, 4810 citations A robust variant of Transformers for Vision
3	Learning Transferable Visual Models From Natural Language Supervision OpenAI, 3204 citations CLIP, image-text pairs at scale to learn joint image-text representations in a self-supervised fashion

Source: Zeta Alpha

Biopharma is in a strong position to adopt AI

In the previous section we showed why AI seems especially well-suited to biopharma. But we think the reverse is also true: that biopharma is in a particularly strong position to adapt to AI.

Biopharma operates a business model unlike any other industry's: The core activity is developing new drugs, which is extremely difficult – and hence slow, expensive and inefficient – but the potential payoff is massive profits for drugs once they are approved. That is until they go off-patent, at which point the profit virtually disappears. All this means there is plenty of scope to put R&D funds into AI, and also a huge scope for AI to bring improvements.

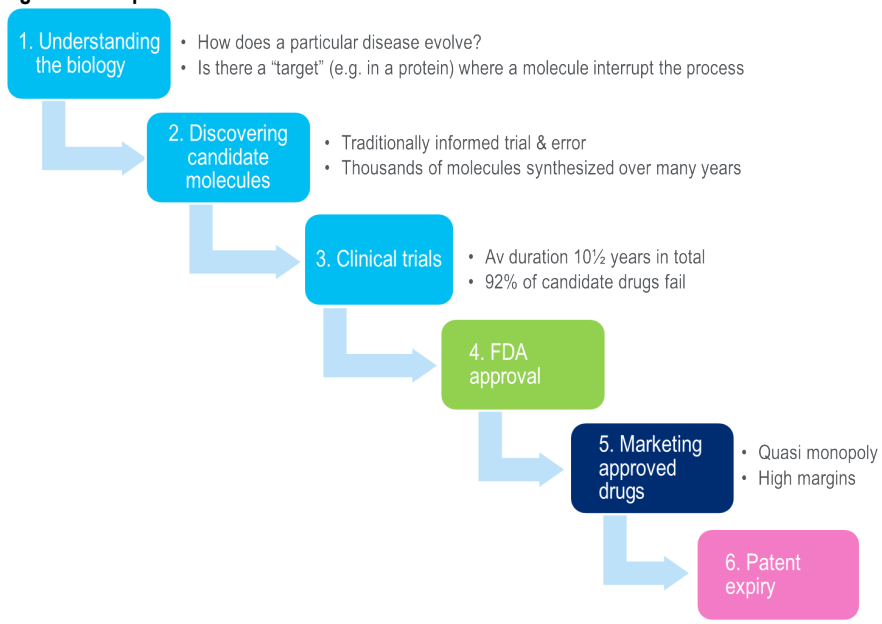
What the contrast with "tech" and consumer staples illustrates

In some ways, the industries that are most like biopharma are consumer staples and "tech" -- consumers staples aren't significantly affected by the economic cycle, and tech evolves through R&D-driven advances. But we think the differences are very illuminating. Tech moves at a speed quite unlike pharma and is getting ever-more efficient. And consumer staples don't have the same life-cycle issues – the very long stage of cash negative development, and the patent cliffs.

Biopharma's unusual economic model

To understand why we think biopharma is in a strong position to adopt AI, it's necessary to understand how the industry works, at least in a simplified way: Figure 10 sets this out.

Figure 10. Biopharma activities



Source: Citi Global Insights, Biomedtracker and Pharmapremia

Stage 1: Finding a target for the drug to work on

Typically occurs in academia

The first stage in developing a new drug involves understanding how a disease or clinical syndrome evolves, and developing a hypothesis on how either suppressing or activating a particular biochemical process will help. The aim is to find a *target* – which might be a protein or another pathway – that researchers want either to interrupt or reinvigorate in order to treat the disease.

This stage – understanding to biology to find a target – usually takes place in academia, and can take decades.

Stages 2 and 3: Creating and testing drugs

Typically conducted by biotechs, at least until Phase III

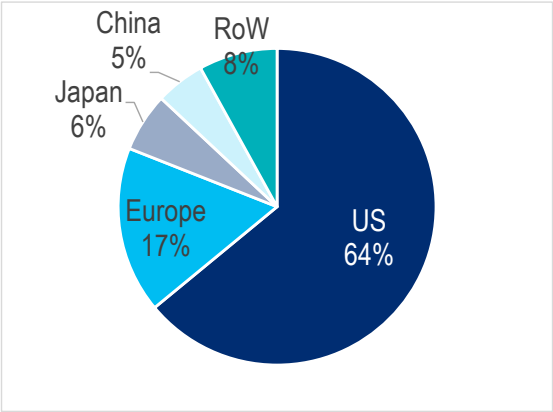
The process of actually creating and testing new drugs is almost always undertaken by biotech firms, and can take something like 15 years and billions of dollars. However, at this point there is no revenue involved, so this process is heavily loss-making.

Stage 4: FDA approval

A drug can be sold only when it receives approval from the FDA, or its equivalents overseas. In practice the FDA is the gatekeeper to the industry globally, because of the economic importance of the U.S. and the high prices paid for prescription drugs there.

Figure 11. Sales of new drugs by region

Drugs launched 2016-21



Source: EPIA and Citi Research

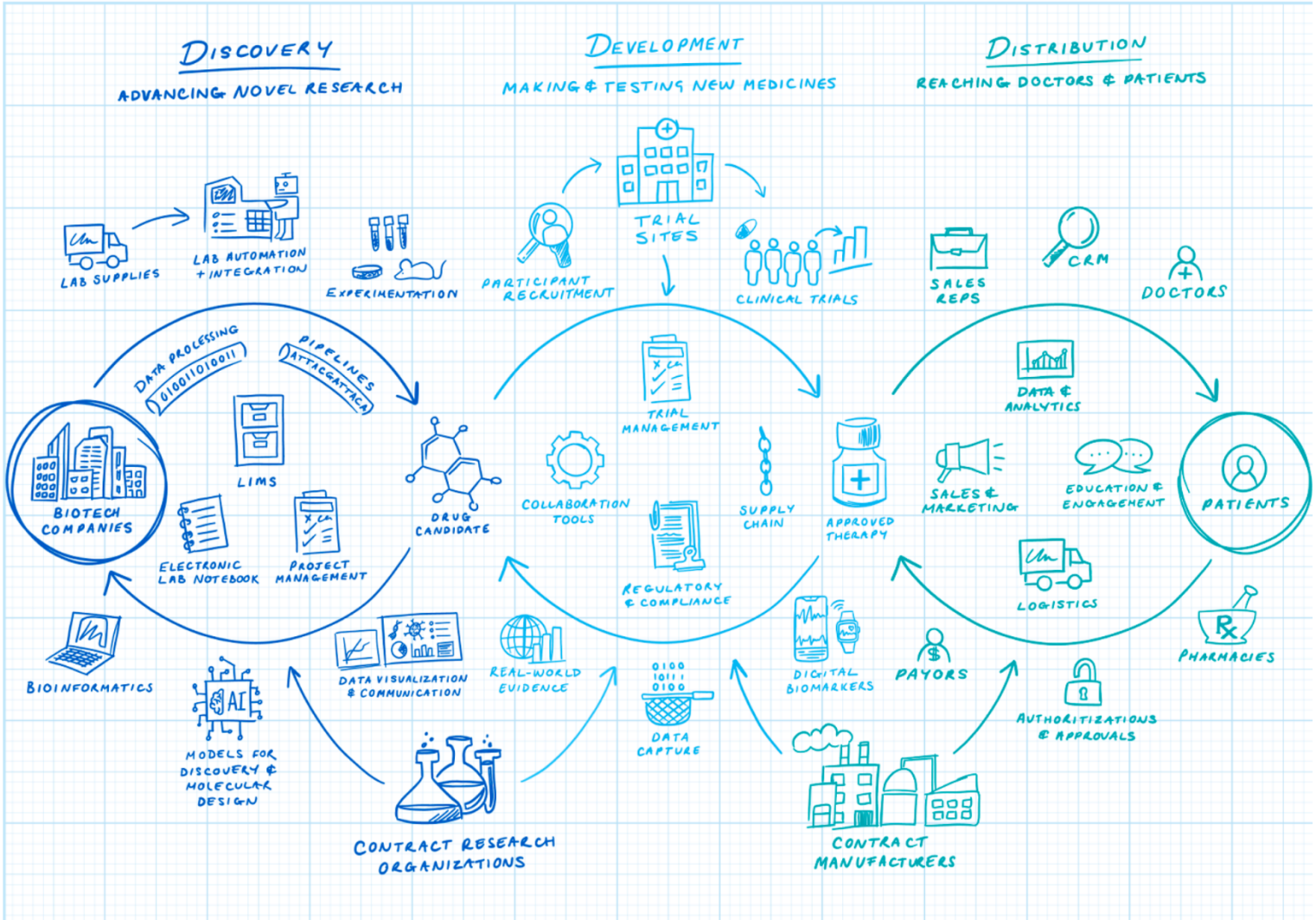
Stage 5: Marketing approved drugs

While drugs are in patent, they are usually extremely profitable – and hence the main pharma companies are both very large in terms of market cap, and have very high margins.

Stage 6: Patent expiry

Once the patent expires, however, generic copies are likely to enter the market, causing the price to drop and the margin to more or less disappear. A typical patent lasts 20 years, and it needs to be registered before the clinical trials, so it may allow the company only 10 years to seek a return on the investment. In practice, there are often ways to extend the patent, for example by introducing new indications (situations in which the drug can be used.)

Figure 12. The Biopharma Industry



Source: a16z Bio + Health

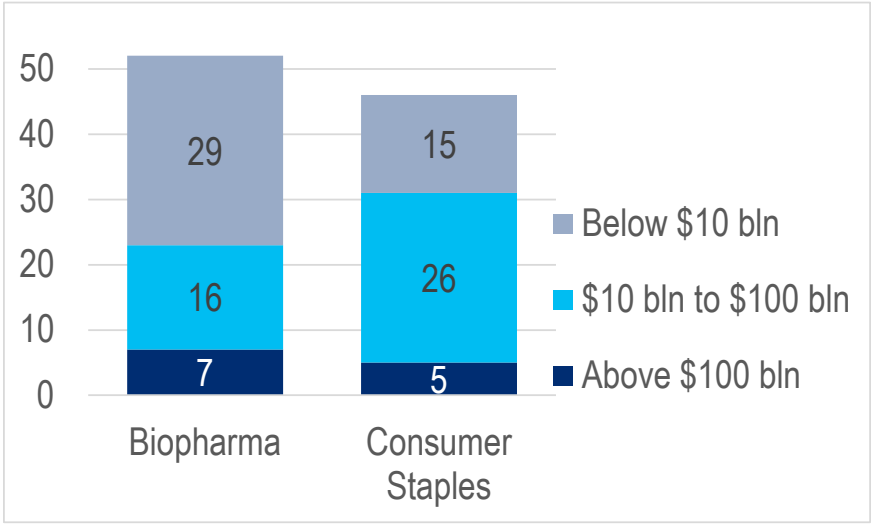
There's a huge range of activities at a granular level

But on top of all this, the granular activities undertaken by biopharma are extremely complicated. Figure 12 on the previous page shows a sketch of all the activities, created by the venture firm A16Z. The authors said that all these activities are suitable targets for SaaS in some way, but we think the sketch also shows how complicated drug development and marketing really is.

The corporate biopharma landscape is quite different from the consumer staples one

In some ways the industry that biopharma is most similar to is consumer staples, because both are globalized, and neither is particularly affected by the economic cycle.

Figure 13. Biopharma vs Consumer Staples -- Number of Stock in Dow Jones U.S. Broad Market Index



Data as of May 17, 2023
Source: FactSet

However, the economic model is completely different.

- There are more biopharma companies with market caps of greater than $100 billion, and also a long tail of companies with market caps of below $10 billion, but fewer in the between. This is illustrated nicely be Figure 13, which shows the number of stocks in the Dow Jones Broad Market Index in the two areas.
- The profitable pharma companies tend to have very high margins. (Of the pharma companies with market caps above $100 billion, the median net margin[15] for FY22 was 38%, vs 18% for the five corresponding consumer staples companies.)

[15] Net income as a percentage of sales.

- By contrast many of the small biotech companies are loss-making. (Of the biotech companies in the Broad Market Index with market caps below $10 billion, 66% were loss-making in FY22. By contrast only 2 of the 15 (=15%) of corresponding consumer staples companies were loss-making.)

And also quite different from “tech”

Another important comparison is with “tech” – and again we are struck by the differences, both in terms of productivity gains, and the time taken to develop new products.

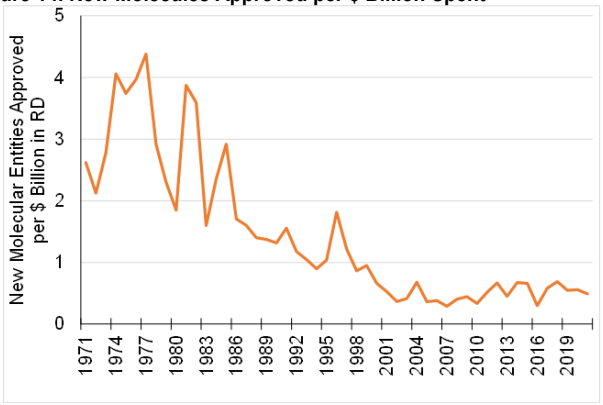
Tech’s productivity increases exponentially whereas biopharma is stuck in low gear

One of the biggest contrasts is in the rate of change of productivity: In biopharma productivity has stagnated whereas Moore’s Law propels tech’s productivity ever-upward.

Figure 14 shows the number of new drugs developed per billion (2021) dollars: there was a significant fall between the 1970s and the early 2000s, and then a very modest increase. Figure 15 illustrates Moore’s law. It isn’t quite comparable, as it shows transistors per chip, not computing power per dollar. Nonetheless the contrast is striking.

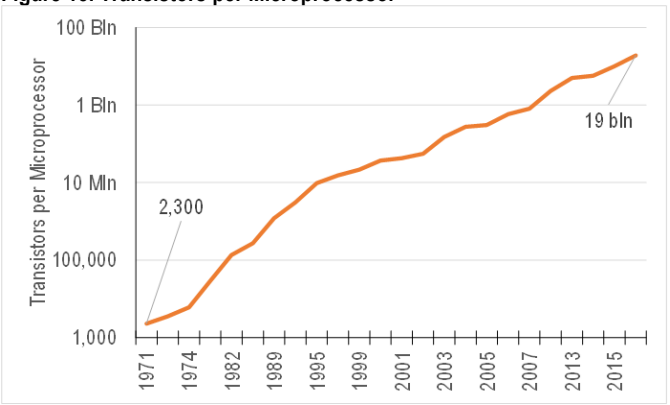
The difference between these two charts provides one of the explanations of why AI can make significant inroads into the biopharma space. Moore’s Law implies AI will get more and more powerful. Figure 14 suggests that financial efficiency in biopharma has barely changed in the last 20 years, despite the innovation in biomedical techniques – e.g. the development of immunology and CAR-T -- and the exponential decline in the cost of gene sequencing.

Figure 14. New Molecules Approved per $ Billion Spent



Note Inflation adjusted (2021) dollars
Source: PhRMA

Figure 15. Transistors per Microprocessor



Source: Karl Rupp, via OWID

In tech, what’s two years old is commoditized

On top of this we see a really important cultural difference between biopharma and tech: Digital products are constantly (and rapidly) evolving. What was cutting edge two years ago is now seen as commoditized. Partly as a result, the tech industry releases products in beta versions, and often uses open source software.

By contrast in healthcare, products need to be shown to be both safe and effective. Furthermore the assumption is that products are fixed, so it is possible to take more than a decade to gather the necessary evidence for the FDA and other regulators.

Conclusion

When we look at biopharma, we see an industry that is both extremely profitable, but also slow and inefficient, with stagnating productivity. Of course the products must be demonstrated to be safe and effective: FDA regulation isn't going away. It is clear to us, however, that there is a huge amount of scope for AI to improve the way the industry works.

AI and Biopharma

In the remainder of the report we examine what AI is already achieving in biopharma

AI and basic biology

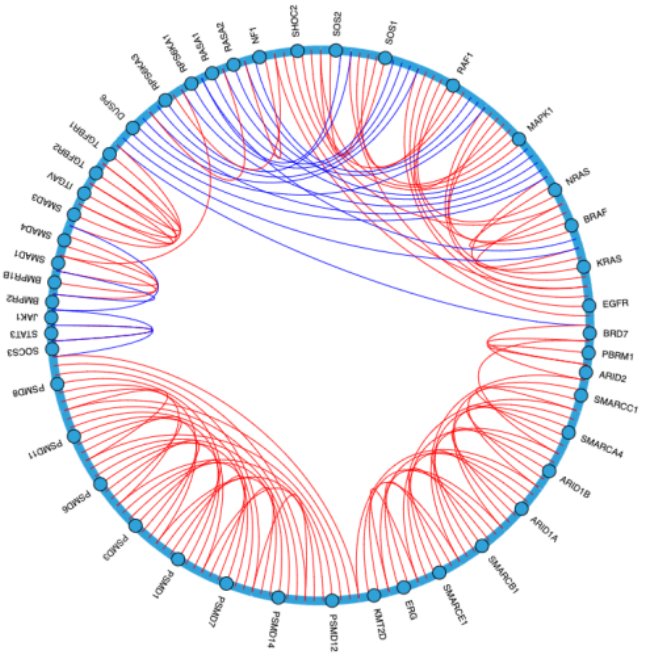
Many AI-first start-ups are focusing on understanding the biology behind particular diseases, the first step in creating an new drug. Generally these companies are creating more data for their AI to work on, usually by automating wet experiments with robotics, which both speeds them up and makes them more precise.

Robotized experiments and machine learning

One example is **insitro**, the start-up founded by Daphne Koller[16]. The company describes itself as "a data-driven drug discovery and development company [that uses] machine learning and data at scale to transform the way that drugs are discovered and developed." The company applies AI both to human cohort data and data from cellular-level experiments, focusing on metabolism and neuroscience. **"**Instead of relying on the limited existing "found" data, we leverage the tools of modern biology to generate high-quality, large data sets optimized for machine learning, allowing us to unleash the full potential of modern computational approaches," the company says.[17]

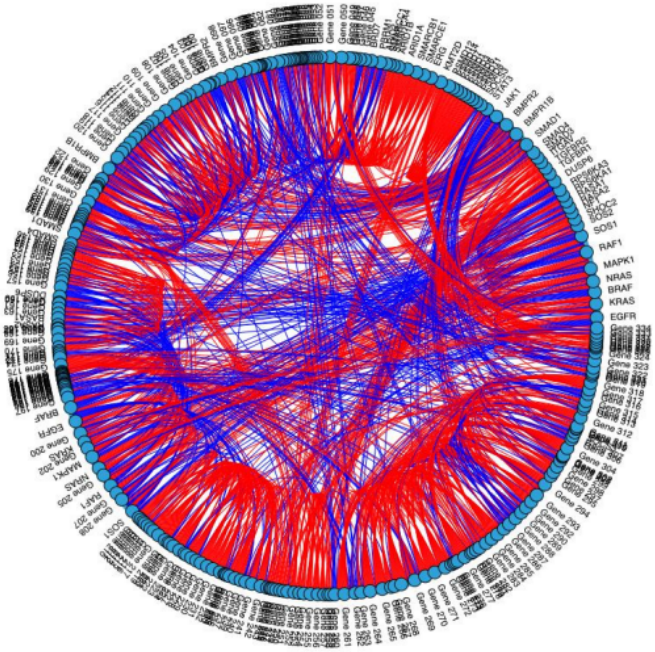
Among other projects, insitro is working with **Gilead** to create experimental models for NASH (a type of liver disease) and with **Bristol Myers Squib** on neuro-degenerative disorders.

Figure 16. Well-known primary relationships between key members of five pathways



JAK/STAT | SWI/SNF | TGFβ | RAS | Proteasome
Source: Recursion

Figure 17. All primary relationships found by Recursion's AI system between key members of the same five pathways



Pathways are: JAK/STAT | SWI/SNF | TGFβ | RAS | Proteasome
Source: Recursion

[16] See page 14 for more details on Daphne Koller.

[17] https://www.insitro.com/about

Another example is **Recursion Pharmaceuticals,** a company based in Utah. It doesn't work only in basic biology, because it has about 50 internally-developed drug programs including five clinical-stage assets. However, its AI systems work (in part) by analyzing tens of billions of relationships between disease models and therapeutic candidates, based in part of its ability to conduct as many as 2.2 million wet lab experiments a week. Recursion's database contains over 3 trillion searchable relationships, which means its data fuels its hypotheses – not the other way around.[18]

The contrast between Figure 16 and 17 illustrates how this works in practice: Figure 16 takes five key protein-related "pathways" (or chains of reactions) and shows how human researchers typically think they are related. Figure 17 on the other hand shows how Recursion's AI sees the relationships between the same five pathways, based in part on its robotized experiments. Clearly this much richer view allows many more processes to be considered.

"Biological software"

Inceptive Nucleics is a small start-up that is trying to take all this further. It also combines highly scalable experiments and deep learning, but its founder and CEO, Jakob Uszkoreit, is aiming to design RNA molecules with entirely new properties.

He says he choose RNA to investigate because it can be applied to so many biological problems, it is relatively simple to analyze (because it uses four different bases only), and it is relatively easy to synthesize.

"RNA seems to be by far the best combination when it comes to the structural simplicity, but also the scalability of synthesis and this experimentation, " he says.[19] "There's huge potential here that so far has been untapped."

He believes this approach will allow Inceptive to create what he calls "biological software" – in other words synthetic RNA that will execute complex functions, specified in a program, within biological systems. The ultimate aim is to be able to design new types of medicines and biotechnologies that are currently "out of reach".

[18] Typically a scientist generates a hypothesis, and then gathers data to test it. With Recursion's approach, data can be gathered and then analyzed without a hypothesis, allowing researchers to find truly novel relationships.

[19] https://bio-eats-world.simplecast.com/episodes/using-ai-to-take-bio-farther

Other approaches

AI can be used to advance fundamental biology is other ways as well.

Apply AI to publicly available data

BenevolentAI's approach is rather different from companies like insitro and Recursion, as its main focus is applying AI to the vast body of publicly-available biomedical data, in order to generate better understanding from it. (It gathers data from over 85 data sources, including from 35 million academic papers, on 'omics, and on biological systems.) BenevolentAI does supplement this with its own experiments, but they're not its most important dataset.

Another point of difference is that Benevolent helps researchers form hypotheses as a starting point for choosing the right target, unlike the companies pursuing hypothesis-free research.

The key aim of Benevolent's AI is to make all the relevant data accessible and useful. It presents the biomedical relationships in a straightforward way to help researchers understand the underlying biology from different angles: some models can rank the entire human genome; others provide promising targets that answer biological questions. Benevolent's AI can therefore guide researchers to ask the right questions and build better hypotheses than may otherwise have been considered, to inform the next steps.

The company does work in drug development. One of their molecules (BEN8744) would be an entirely new type of treatment for ulcerative colitis – and this should enter clinical trials later this year. Benevolent is also developing molecules in other disease areas, with two molecules (BEN28010 and BEN34712) in the pre-clinical stage, and two others that are still being optimized.

BenevolentAI has one molecule (BEN2293) that it took to Phase IIa trials. Unfortunately however the results were less encouraging than it had hoped and in late May it announced it would not invest further in that molecule.

Benevolent has had a partnership with **AstraZeneca**, which was expanded at the beginning of last year into two further disease areas. Benevolent has so far helped AstraZeneca select five new targets for its portfolio.

AI enhanced electron microscopy

Gandeeva Therapeutics is another company offering AI as a service in basic biology, but it uses AI for cryogenic electron microscopy (or cryo-EM). This technique lets researchers see what molecules actually look like almost on an atom-by-atom basis. It works by shooting electron beams at a flash-frozen sample of the molecule of interest, and using AI to reconstruct 3D images of the molecule. It shows the structure of molecules down to about 2 Ångströms, which is only a bit larger than an atom. (An Ångström is 10^{-10} meters. A covalently-bonded carbon atom is about 1.5Å in diameter.) This is important because understanding how proteins react with other molecules often benefits from a combination of theoretical modeling (from the likes of AlphaFold) and observation.

The company says "by imaging our targets at atomic resolution throughout the design process, we steer away from discovery dead ends, and towards progressively higher confidence therapeutics."

AI and drug development

AI is equally important in the drug development phase. This is the stage after a *target* has been found, when biopharma companies develop a candidate drug to take to clinical trials. To do this, they must:

- **Validate the *target***, ensuring it is *druggable* – in other words that (in principle) molecules could bind to it, and if they do there will be a biological response that can be measured. The process of validation results in greater confidence that binding a molecule to the target will result in the desired therapeutic effect.
- **Generate *hit* molecules.** These are the molecules that bind to the target – and typically biotech companies generate thousands of molecules to try to find one that works. This isn't literally trial and error of random molecules, because there are in theory more than 10^{60} molecules that could be formulated, which is far, far too many to try out; rather it is more like "informed trial and error".
- **Optimize the *lead*.** Once a hit molecule has been found, it's necessary to optimize it, to improve potency and selectivity. A candidate molecule has to have the right pharmacokinetic (PK) properties too – it's desirable to have a drug that can pass through the gut, and needs to be taken only once or twice a day.
- **Complete the preclinical studies**. Biopharma companies have to provide evidence that the potential drug is likely to be safe and effective in order to move on to clinical trials. Clinical trials can begin only when the FDA gives it IND[20] status.

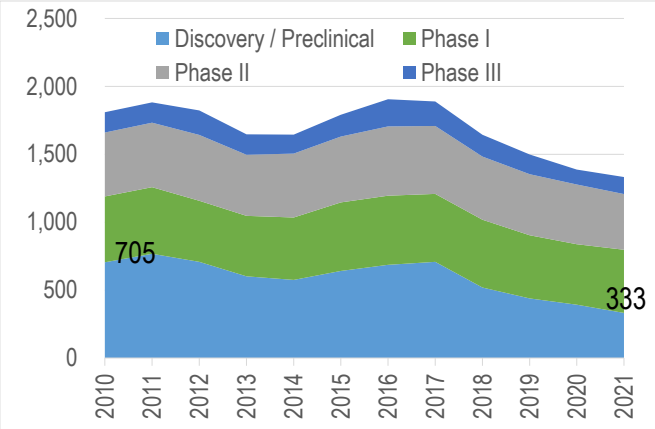
AI is already making a big impact

The impact AI is having in the pre-clinical stages can be seen by comparing Figure 18 and 19.

Figure 18 shows that since 2010 the number of molecules developed by the top 20 pharma companies that have in the preclinical stages has halved, from 705 to 333. (The number of assets in clinical trials has also dropped, albeit by a much smaller percentage, from about 1,100 in 2010 to about 1,000 in 2021.) Figure 19 shows that by contrast, the number of early stage assets generated by AI-first companies has surged in the last five years, and is now equivalent to about half the total generated by the top 20 pharma megacaps.

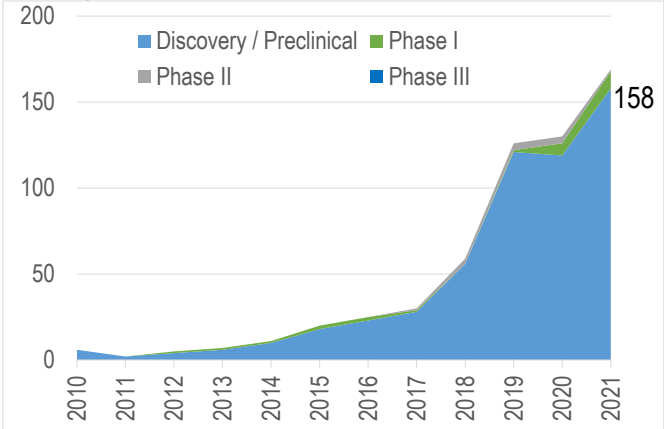
[20] IND: Investigational New Drug

Figure 18. Top 20 Pharma Companies -- No. of Assets in Discovery and Clinical Trials



Source: Nature Reviews Drug Discovery via Recursion

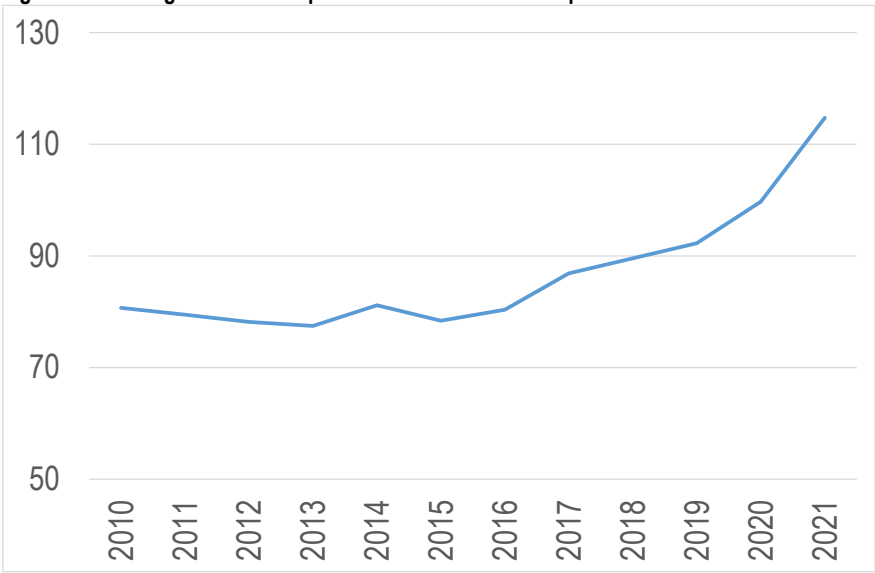
Figure 19. Top 20 AI-First Pharma Companies -- No .of Assets in Discovery and Clinical Trials



Source: Nature Reviews Drug Discovery via Recursion

The decline in the pharma companies' preclinical pipeline can be partially explained by them becoming more selective. However it also comes against a backdrop of increasing spend on R&D, as Figure 20 shows.

Figure 20. Leading Pharma Companies -- Combined R&D Expense in Billions of Dollars



Data captures R&D expense from 12 of the leading pharma companies globally.
Source: Company Reports

What AI can deliver

The power of AI in this stage is shown by some statistics from **Exscientia**, one of the full-stack AI-first biopharma companies. The company says its AI means it can:

- Reduce the discovery time by 70%, relative to conventional processes;
- Reduce the number of compounds that need to be synthesized by 10x; and
- Reduce the capital required to go from finding a biological target to taking a molecule through preclinical toxicological studies by 80%.

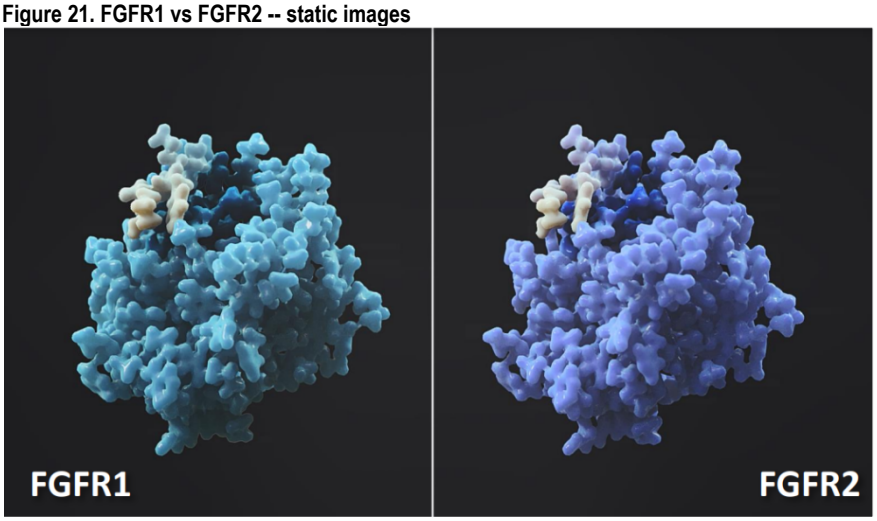
AI and protein movement

Relay Therapeutics offers an example of how AI can help identify hits and optimize the leads, in oncology and hereditary diseases. It specializes in understanding protein motion to guide small molecule drug development with the aim of “drugging previously undruggable targets”.

The proteins in our body are never static – they are moving all the time and their functions are highly dependent on their motions. There are many proteins that have nearly identical structures and perform similar yet different functions because they move differently. Structural similarities in proteins from the same family make it hard to find a hit molecule for a specific target without causing serious side effects on other proteins in the family.

One example is the FGFR (Fibroblast Growth Factor Receptor) family. This consists of four very similar proteins that are crucial in the normal development of animal cells, but are also involved in several types of cancer. Inhibiting the activity of FGFR 2 has been proven to be an effective cure for a certain type of bile duct cancer, for example, but therapies based on the existing non-selective pan-FGFR inhibitor are constrained by a dose-limiting side effects caused by inhibition of FGFR1 and FGFR4.

Figure 21. FGFR1 vs FGFR2 -- static images

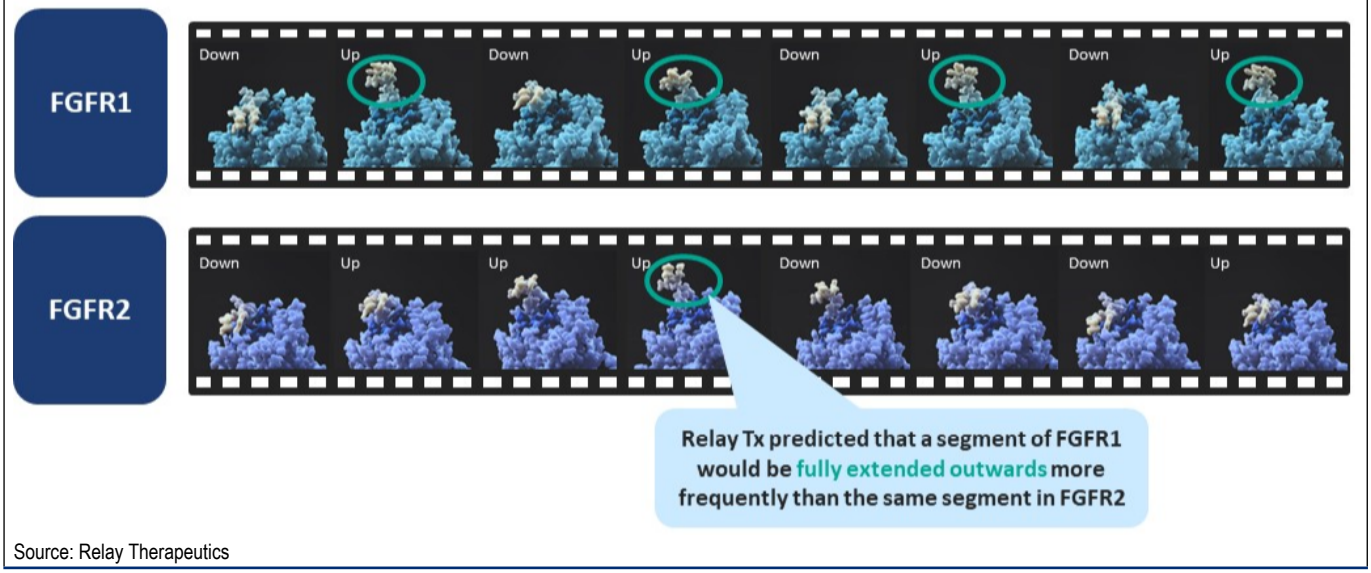


Source: Relay Therapeutics

Relay is able to use mathematical biophysics to predict how proteins move, calculating how the position of each atom varies every femtosecond. (A femtosecond is 10^{-15} of a second.)

In turns out, for example, that there is a section of the FGFR1 protein that is fully extended more frequently than the corresponding section of FGFR2, as Figure 22 shows. Relay has been able to use this to develop a new potential drug (RLY-4008) which inhibits FGFR2 but that isn't limited by creating off-target toxicities of high phosphate level in the blood (related to FGFR1) and diarrhea (related to FGFR4).

Figure 22. FGFR1 vs FGFR2 analyzed with Relay's motion-based approach



Source: Relay Therapeutics

RLY-4008 is now in a Phase I/II trial. Crucially Relay has chosen the participants to ensure they have cancer that has been altered by FGFR2 and as such are more likely to respond well to a drug – like RKY-4008 – that inhibits that particular protein. In effect this means the trial is more likely to be successful -- but it also means that -- assuming the drug is approved -- it will be available only for patients with FGFR2-modified tumors.

Relay's approach

RLY-4008 illustrates Relay's overall approach:

1. The company starts with a target that is well understood from a biological point of view, but where in practice it has proved impossible to find an effective drug.

2. It then works out if it can find an area of the protein where a small molecule could bind, based on Relay's understanding of the protein's movements.

3. Relay then simulates how millions and millions of small molecules would interact with the target, given the protein's movement, in order to find one that is likely to bind effectively. This is more akin to brute force than design.

4. The company then carefully optimizes the molecule to ensure that it has best possible qualities, for example around pharmacokinetics.

Currently, Relay Therapeutics is working on more than 10 drug programs, with four molecules in clinical trials, one of which is RLY-4008.

Using machine learning to choose antibodies

AbCellera Biologics works only in developing drugs – in this case antibodies. It doesn't attempt to find targets, and it leaves others to actually conduct the clinical trials.

The company aims to be the partner of choice in the field. Currently it is working with over 40 different partners -- megacap pharma companies, small biotechs and non-profits -- on more than 175 programs. The partners come to AbCellera with an idea for a new antibody-based drug and a specific target they want to address; AbCellera's technology helps find the best antibodies for the target, allowing the partner to develop them further and prepare them for clinical trials.

It's very much an AI-first biopharma company, but it works in a way that's a bit different to many of the others. It doesn't attempt to "design" drugs; rather it uses machine learning to pick what it believes is the best antibody from the millions the body produces, in order to address the target.

To do this, it generates, stores and processes huge amounts of data, and this in turn requires the most up-to-date soft- and hard-ware, and also roboticized wet-lab experiments. (Whereas BenevolentAI aims to integrate data from many different sources, AbCellera doesn't rely on third-party data.)

The result is carefully-curated data that allows researchers to quickly explore and interpret complex information relating to antibodies, which in turns lets the company select the optimum antibodies for the job.

The first antibody for Covid

The way it develops antibodies is illustrated by the story of bamlanivimab, the first monoclonal antibody therapy for covid that was authorized by the FDA. Bamlanivimab was developed in partnership with **Eli Lilly**.

Back in March 2020, right at the start of the pandemic, AbCellera received a blood sample of one of the first Americans who had recovered covid. The company knew that that sample would include antibodies that had successfully overcome the virus; the question was finding them.

AbCellera had already developed a device, about the size of a credit card, with over 200,000 chambers on it. It placed different cells from the blood sample in each of these chambers – and each of those cells represented a different antibody. It then added the covid spike protein, and a chemical reagent that glowed green if the antibody cell attached successfully to the spike. It found the relevant chambers using an imaging device that could focus closely enough to actually see which turned green.

In fact AbCellera assessed about 5½ million potential antibodies from the original blood sample within 3 days – and by day nine, it had sequenced the DNA of 500 antibodies that had successfully attached to the spike.

The next question was working out which of these would be best to use as a real drug. AbCellera assessed these 500 candidates, according to 500 different parameters, effectively creating a spreadsheet with 250,000 cells that needed to be filled in – far too many for a human to consider.

By rapidly reproducing and testing the antibodies, AbCellera – working with Lily and the U.S. government Vaccine Research Center -- was able by day 23 to filter the group down to 24 antibody front-runners. These were passed to Lilly for further development.

The first human was injected with candidate antibodies in the Phase I trial only 90 days after AbCellera received the blood sample. Bamlanivimab was given emergency approval by the FDA in November 2020.

T-Cell Engagers

Since that time, AbCellera has further developed its techniques, and it is now focusing more on antibodies that have previously eluded other researchers. One example of that is the focus the company is currently putting on a class of antibodies called T-Cell engagers, or TCEs. These have the potential to provide powerful therapeutics on for a wide range of cancers.

TCEs combine two highly targeted arms: One arm that binds with T-cells (the cells the body creates naturally to fight cancers) in a region called CD3; the other arm bind directly to the cancerous cells. The aim is to bring the T-cells and the tumors close together by binding to both targets at the same time. To be successful, a TCE needs both to activate the right sort of T-cell in right way; and to bind specifically to the targeted tumor. Until now, the main problem has been the limited number of molecules that engage with the T-cells appropriately.

Figure 23. T-cell engager

Source: AbCellera

AbCellera has used its ML techniques to find many hundreds of different arms, both to fine tune the activation of the right T-cells, and to bind with cells of specific types of tumor. The company is increasingly confident its AI will allow it to create custom-designed TCEs for a wide range of cancers.[21]

[21] Previously AbCellera used ML to select existing antibodies – it didn't actually "make" or "design" anything new. TCEs are a little different, each arm is a pre-existing molecule, but because AbCellera slots them together, it is creating entirely new molecules.

“In late 2021, we recognized a gap that was preventing powerful TCE cancer treatments from making it to patients and felt confident that our antibody discovery and development engine could provide the solution,” said Murray McCutcheon, who is in charge of partnering at AbCellera. “In 18 months, we have developed a TCE discovery platform and are leveraging the extensive datasets we’ve generated to custom-build TCEs and help bring better cancer treatments to patients faster.”

AI and clinical trials

The clinical trials represent the longest, and most expensive, stage in bringing a drug to market. We believe AI can help make them more efficient, although its impact is likely to be less dramatic than in the earlier stages of drug development.

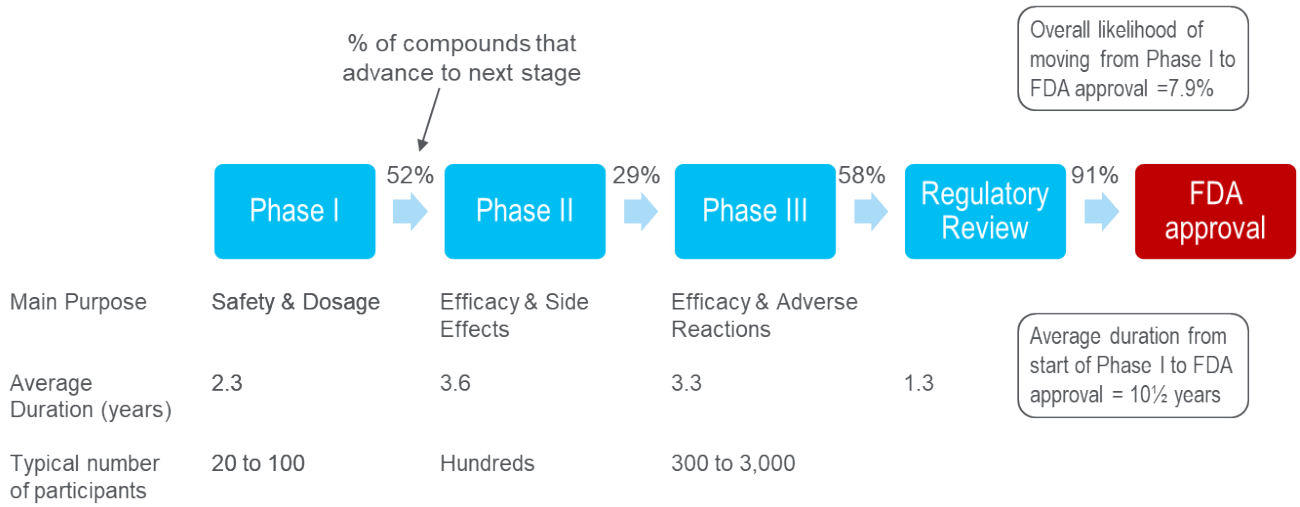
Understanding clinical trials

Clinical trials are necessary to show a drug it is effective and safe. There were just under 10,000 Phase I, II or III trials registered with the FDA in between 2010 and 2020, and Figure 24 shows the average of how they moved through the system. On average:

- The likelihood of approval (or LOA) – in other words moving from Phase I to FDA approval – was about 8%. The biggest fall-out occurs after Phase II: only 29% of drugs in Phase II moved to Phase III.
- On average it took 10½ years to move from the start of Phase I to FDA approval.

One way of looking at these two stats – that only 8% of molecules made it through the trials, and that on average they take 10½ years – is to conclude how difficult the drug discovery process is. It's also possible to look at them and conclude that the process is highly inefficient.

Figure 24. Clinical Trials



Data based on all 9,704 FDA-registered clinical trials in the Biomedtracker database between Jan 2010 and Nov 2020.
Source: Biomedtracker and Pharmapremia

More detailed numbers from the Biodmedtrack database show that the likelihood of approval is correlated with development timelines: disease therapies and types of approach with better LOAs tend to be approved quicker. Furthermore where the scientists have a really detailed knowledge of the biology and therapy, the clinical trials seem to go better.

- Rare disease therapies have much-better-than average likelihood of approval (17%), whereas therapies for chronic, high-prevalence diseases have a worse-than-average LOA (6%). (See Figure 25.)

- Biological complexity in modalities – in other words the approach taken – tends to lead to higher likelihoods of approval. Figure 25 shows CAR-T and RNA interference have had the best LOAs.
- Development programs where the trials preselected participants using biomarkers also had comparatively good likelihoods of approval. (16%)

Figure 25. Likelihood of approval by type of disease, drug modality and patient selection

	Likelihood of Approval	# Drugs	% of Drugs
Type of Disease			
Rare Disease	17%	1,256	10%
Chronic, High Prevalence Disease	6%	1,978	16%
Other	7%	9,494	75%
Treatment Modality			
CAR-T	17%	67	1%
siRNA/RNAi	14%	87	1%
Monoclonal Antibody	12%	2,136	17%
ADCs	11%	184	1%
Gene therapy	10%	96	1%
Vaccine	10%	316	2%
Protein	9%	800	6%
Peptide	8%	619	5%
Small molecule	8%	7,171	56%
Antisense	5%	162	1%
Patient Preselection			
Preselection Biomarkers	16%	767	6%
No preselection biomarkers	8%	11,961	94%
All Diseases	**8%**	**12,728**	**100%**

Source: BioMedtracker

Figure 26. Likelihood of approval by disease area

	Likelihood of Approval	# Drugs	% of Drugs
Hematology	24%	352	3%
Metabolic	16%	399	3%
Infectious disease	13%	1,170	9%
Others	13%	541	4%
Ophthalmology	12%	415	3%
Autoimmune	11%	1,305	10%
Allergy	10%	201	2%
Gastroenterology	8%	186	1%
Respiratory	8%	501	4%
Psychiatry	7%	442	3%
Endocrine	7%	887	7%
Neurology	6%	1,411	11%
Oncology	5%	4,179	33%
Cardiovascular	5%	651	5%
Urology	4%	88	1%
All Diseases	**8%**	**12,728**	**100%**

Source: BioMedtracker

AI can accelerate clinical trials

We believe AI will be able to improve the likelihood of approval, and shorten the trials, in a variety of ways.

Ben Taylor, the CFO and chief strategy officer at Exscientia, says that if the molecule is designed rather than discovered, that changes the nature of the Phase 1 and 2 from *"exploring properties"* to *"confirming design".*

AI can also help in clinical trials in at least two further ways:

- **Finding the right patients for the trial**: In some diseases, different patients respond to different drugs. (In certain types of cancer as few as 1 in 10 patients respond to the therapy chosen by their oncologist.) With AI it's possible to measure how live cells from an individual patient respond to the available drugs at the single cell level, and then select the best option for that patient. In clinical trials, a potential drug can be tested only on patients who are likely to respond, greatly increasing the chances of a positive outcome.

- **Finding the right dosing**. One of the main purposes of Phase 1 and 2 trials is to find the optimum dose for a candidate drug. AI makes it possible to run *in silico* trials to estimate both the pharmacokinetics (in other words how the drug interacts with the body) and the pharmacodynamics (in other words what effect the drug has). The optimum dose has to be confirmed from clinical trials, but the entire process is faster and easier if one starts with a good estimate. **Certara** is a global leader in this type of *in silico* trial. It also provides AI toolkits to help researchers interpret clinical trial data helping drug developers make better-informed decisions regarding formulation, dosing, side effects and regulatory reporting.

As discussed on page 6, Exscientia's EXS21546 is an anti-cancer drug that works by blocking the effect of adenosine of certain tumor cells, because adenosine suppresses the body's immune response to the cancer. But this approach is likely to work only in patients with tumors that have high concentrations of adenosine. Until now, however, there hasn't been a good way of identifying them. This means that all previous trials included many patients for whom the drug was unsuitable.

Exscientia, however, has created an adenosine burden score by using an AI-driven analysis of RNA levels in single cells – and this should allow it find the right participants and therefore increase the likelihood of a successful outcome.

Vida Lung is an example of an AI-driven company that works in partnership with the biopharma industry, helping accelerate trials by applying AI-driven image recognition analysis to X-rays and CT-scans of lungs.

"Clinical trials have traditionally been slow and inefficient, which is why they're so expensive," Susan Wood, the CEO, told us. "If we can make them more efficient – by finding the right patients and sites quickly and measuring treatment response more precisely – that creates significant value."

Vida says that its systems help make clinical trials more efficient in several ways.

1. **By finding participants.** Trials are expensive for many reasons, and one of them is finding an appropriate (and demographically diverse) cohort of participants, especially if the therapy is aimed at highly specific type of patient. If a person has a routine lung scan, Vida's AI can automatically analyze the image and identify individuals who meet enrolment criteria. This analysis allows the attending clinician to match that the person with a suitable trial. Vida has a network of more than 1,000 sites, allowing trial sponsors to quickly reach populations in many geographies.

2. **By leveraging digital biomarkers.** By analyzing high resolution images deeply, Vida's AI can precisely measure changes in structure and function over time, often well before these changes will become apparent in conventional clinical tests (for example pulmonary function tests). Figure 25 shows that clinical trials that select participants are more likely to succeed than those that don't.

3. **By enabling better imaging data workflows.** Due to its cloud-based platform, which automates tasks for participating sites, Vida leverages AI for quality control, and simplifies the operations of an imaging-based trial.

"The use of AI resonates with biopharma companies because they operate in a high stakes environment of expensive clinical trials with high failure rates," Dr. Wood said. "They need the increased quality of information and efficiency that AI provides."

Summary

This is the fourth report in our series *Smart Thinking on AI in Healthcare.* The conclusions are summarized in Figure 27.

We believe both that (1) AI is especially important for biopharma, because AI can understand biology in a way humans cannot, and that (2) biopharma is in a particularly strong position to adopt AI, because it is simultaneously highly profitable and extremely inefficient.

Drug development

The reason it cost so much, and takes so long, to develop new drugs, is that biology is incredibly difficult. The industry traditionally talks about “drug discovery”, and it involve a lot of iterative trial and error, with researchers slowly closing in on candidate drugs. AI has the power to change that, because it promises to radically improve our understanding of biology and as a result developing new drugs will be much closer to *design* than *discovery*.

Clinical trials

Clinical trials will always be necessary to show drugs are both effective and safe. AI, however, has the power both to speed them up and to increase the likelihood of success.

- AI-first drug development alters the nature of Phase I and II trials, making them more about *confirming* that a drug works as it is designed, and less about *exploring* its properties.
- AI can be used to find patients who are more likely to respond well to the drug.
- AI can help find more participants more quickly, including from historical under-represented populations.

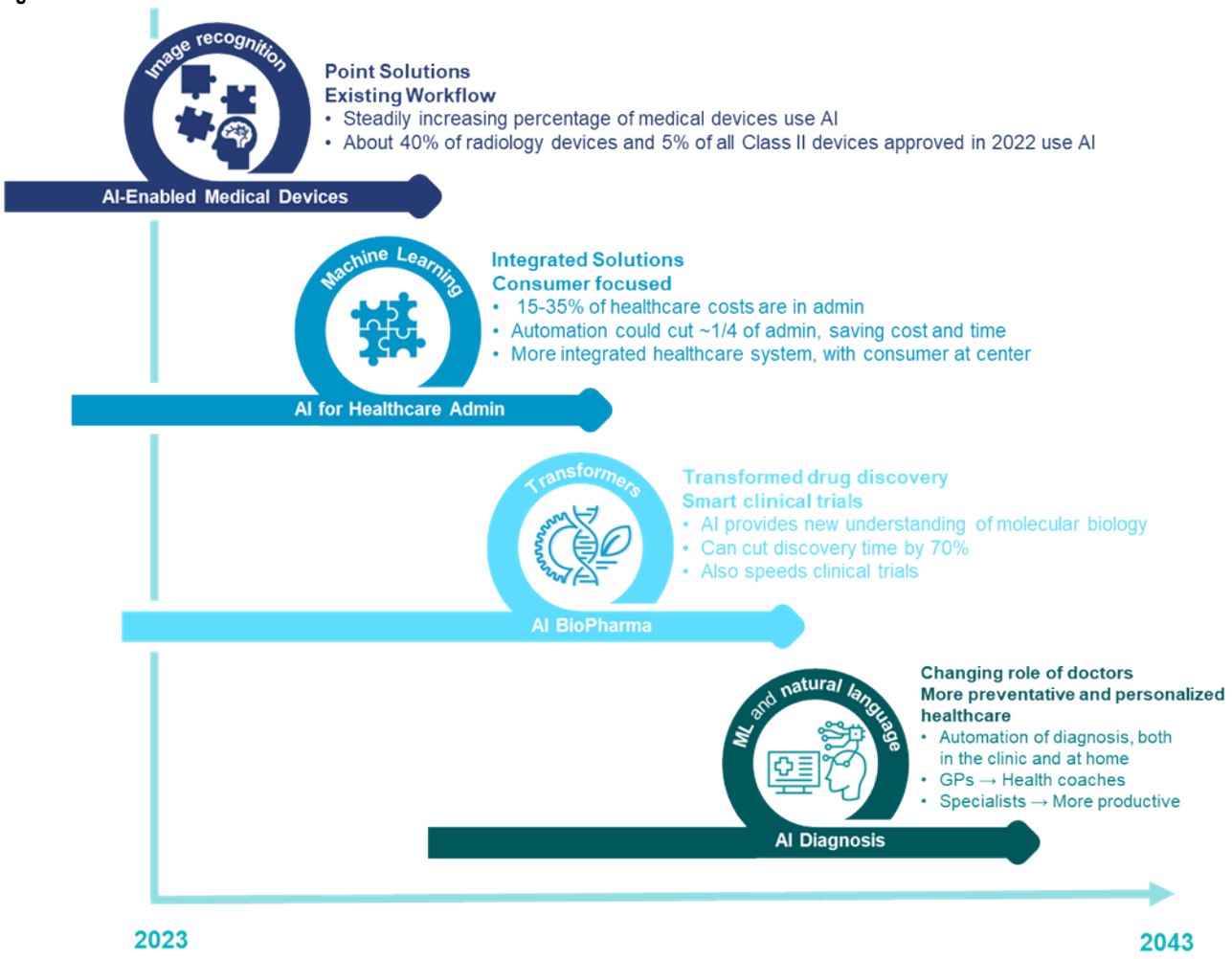
Use cases

This report has discussed how AI-first biopharma companies like AbCellera and Relay have used AI to develop new drugs and how companies, for example Exscientia and Vida Lung, are using AI to improve the chances of clinical trials succeeding.

Speed of change

In summary we expect AI to have a profound effect on the drug industry. Companies that are able to adapt to it – that can combine life science and data science effectively -- are likely to thrive. However it’s important to remember that biopharma moves slowly, and that drugs will always need to be proved to be efficacious and safe in clinical trials in humans.

Figure 27. AI in Healthcare



Source: Citi Global Insights

If you are visually impaired and would like to speak to a Citi representative regarding the details of the graphics in this document, please call USA 1-888-800-5008 (TTY: 711), from outside the US +1-210-677-3788

Disclosures