LIFE SCIENCES SNAPSHOT

A Quarterly Report on Financing Trends

AI IN DRUG DISCOVERY
Q3 2023

orrick

Data provided by

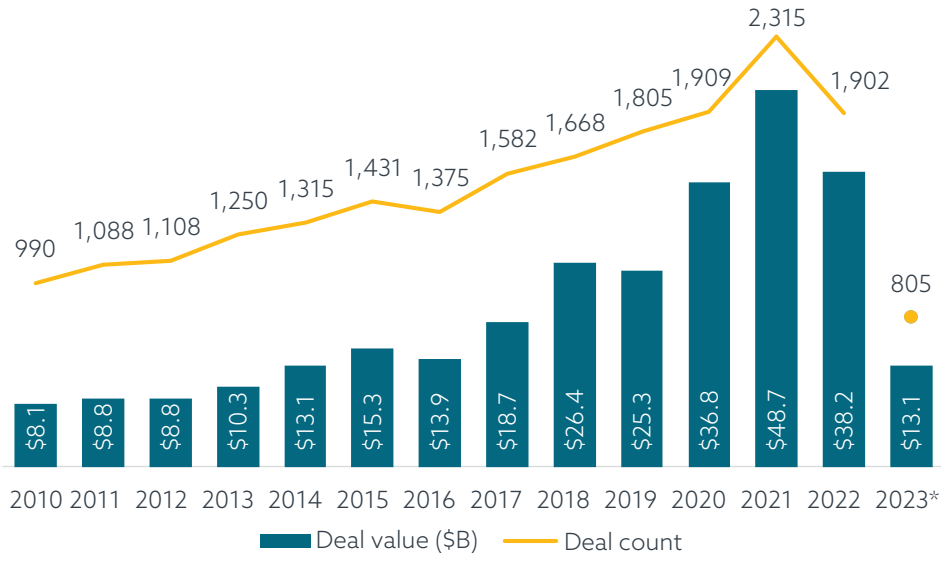
PitchBook

Key Takeaways

This report series examines quarterly trends in life sciences venture investment. Key findings for Q2 2023 include:

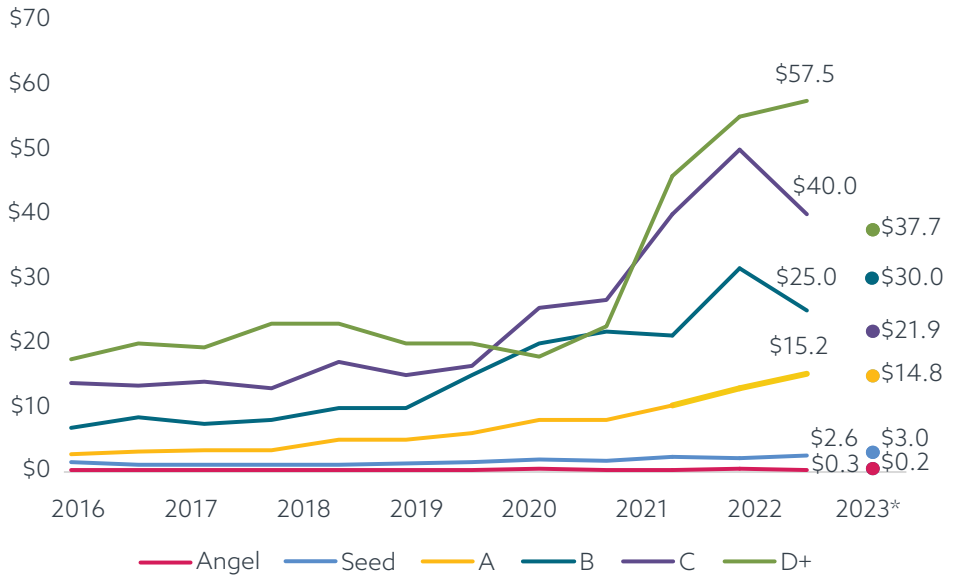
- Life sciences VC deal value in Q2 2023 totaled $6.7 billion across 384 deals, representing a slight increase in value and a moderate decrease in deal count from the previous quarter. YTD deal value of $13.1 billion represents 34.3% of 2022's annual total.
- Median deal sizes declined moderately for most deal series YTD, including a material 34.4% decline for the latest category of Series D+.
- Median pre-money valuations plateaued YTD for the angel, seed, and late-stage VC categories, while the median valuation for the early-stage VC category (through Series B) grew 18.8% YoY—hitting a record high of $49.9 million.
- Total exit value rose to $2.0 billion in Q2 from $1.4 billion in Q1, but activity remains historically low. Deal value halfway through 2023 represents 19.4% of 2022's annual total.

Life sciences VC deal activity



Source: PitchBook | Geography: US
*As of June 30, 2023

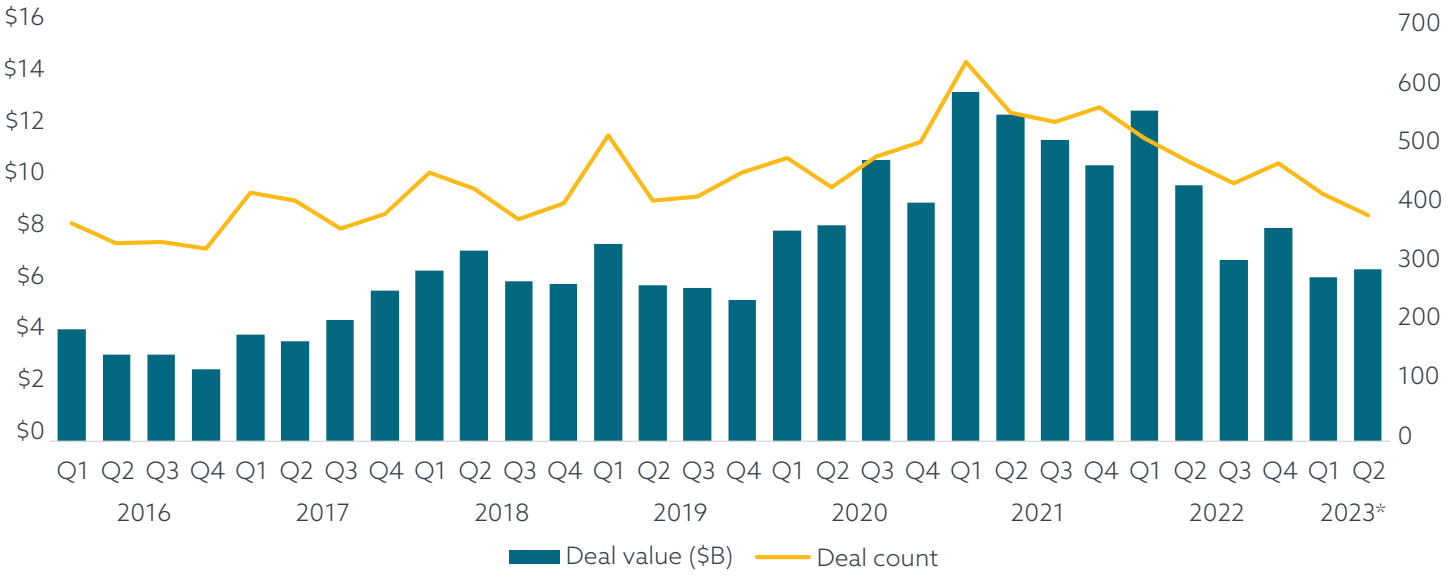
Median life sciences VC deal value ($M) by series



Source: PitchBook | Geography: US
*As of June 30, 2023

Market Analysis

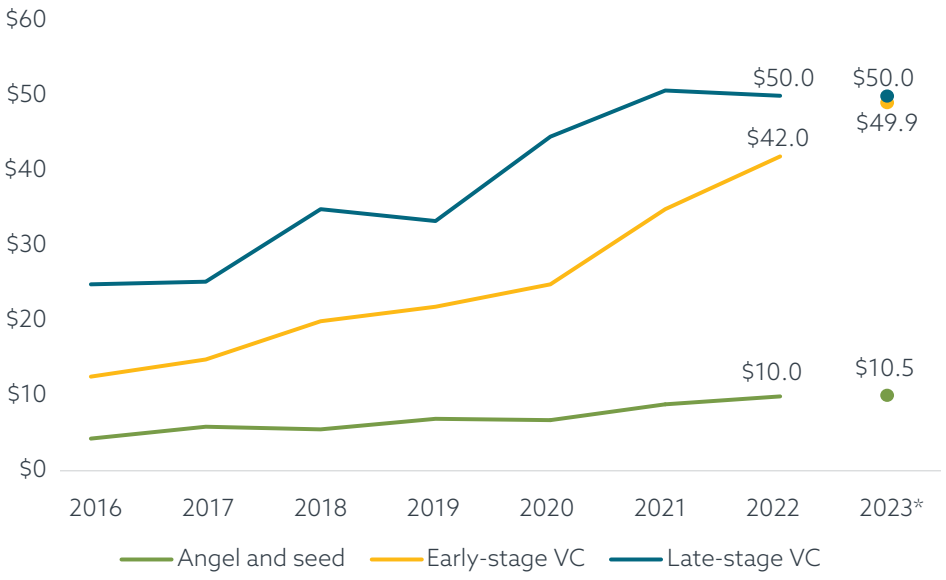
Life sciences VC deal activity by quarter



Source: PitchBook | Geography: US
*As of June 30, 2023

2023 began on a low note for the broader VC industry, and difficulties remain. Deal activity in Q2 2023 totaled $6.7 billion across 384 deals, representing a slight increase in value and a moderate decrease in deal count from the previous quarter. This reflects the recommitment of investors to a more selective population of portfolio companies and the flight to quality that LPs are seeking amidst rockier market conditions. YTD deal value of $13.1 billion represents 34% of 2022's annual total, highlighting the continued impacts that macroeconomic factors—including stubbornly high inflation—are having on dealmaking. Another way these factors have manifested is in median deal sizes, which have declined moderately for most company stages YTD except for seed and Series B rounds. Just as in Q1 2023, the latest categories for Series C and Series D+ deals exhibited the starkest declines

Median life sciences pre-money valuations ($M) by stage



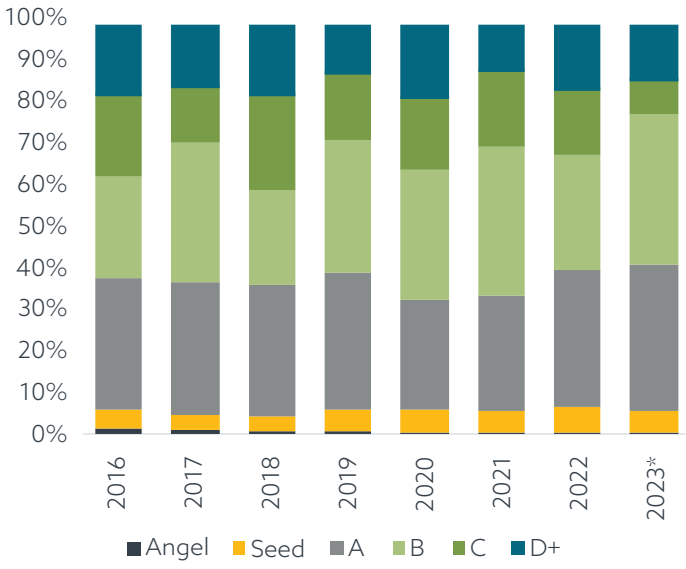
Source: PitchBook | Geography: US
As of June 30, 2023

of 45.3% and 34.4%, respectively. Megadeals are still closing, supported by significant capital reserves held by large funds and experienced managers, but the median is reflective of a more common scenario in which late-stage startups are no longer able to justify the inflated check sizes they received during the VC frenzy that occurred between 2020 and the first quarter of 2022.

Pre-money valuations were once again more resilient than deal sizes, with the median remaining flat for the angel, seed, and late-stage VC categories YTD. The median valuation for the early-stage VC category (through Series B) grew 18.8% YoY, recovering from the declines experienced in Q1 and hitting a record high of $49.9 million. The portion of total deal value driven by early-stage deals has ticked up over the past decade. Late-stage series (C and D+) account for a slightly smaller portion of total deal value and count YTD compared with 2022. In a similar vein, the largest category of individual deals ($100 million and above) accounts for a smaller portion of total deal value when broken out by deal size buckets, though this category has gradually increased its annual portion over the past decade.

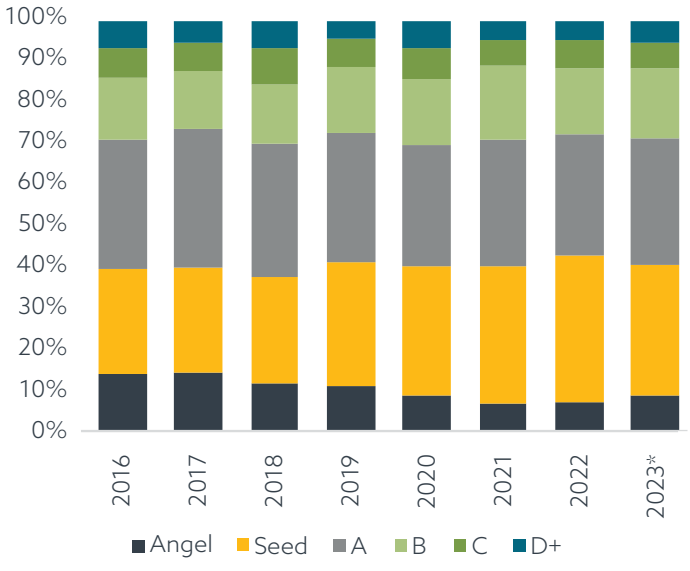
Total exit value rose to $2.0 billion in Q2 from $1.4 billion in Q1, but activity remains historically low and YTD value halfway through the year represents 19% of 2022's annual total. IPOs represent more than half of total deal value YTD. Exit activity has not yet recovered to pre-COVID-19 levels, with just $3.4 billion closed across 51 deals YTD—though, there are some signs of life. Seven life sciences companies have gone public via IPO, driving more than half of YTD exit value, despite the prevailing narrative that the IPO environment is unsound. Acquisitions account for the largest portion of exit count with 32 closed YTD. Life sciences companies hold a degree of insulation from valuation drops and investor skittishness due to their unique milestones and timelines to market, particularly for therapeutics. Progression through the clinical trial and drug approval process remains a reliable way for companies to secure significant returns through Big Pharma acquisitions.

Share of life sciences VC deal value by series



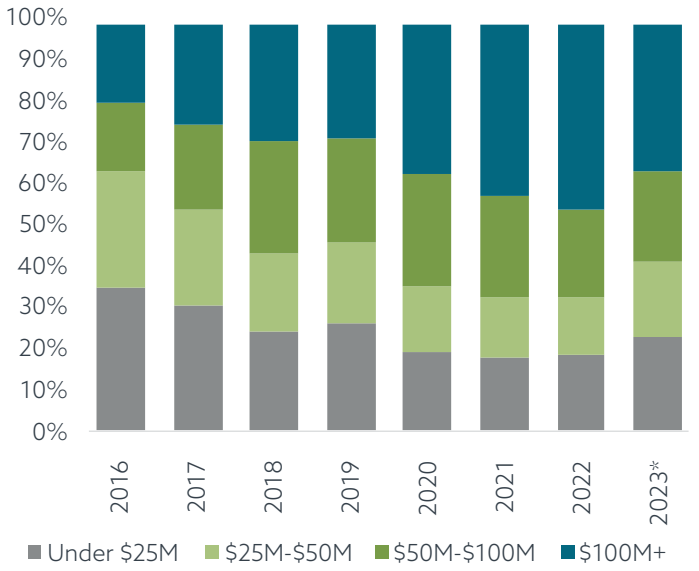
Source: PitchBook | Geography: US
*As of June 30, 2023

Share of life sciences VC deal count by series



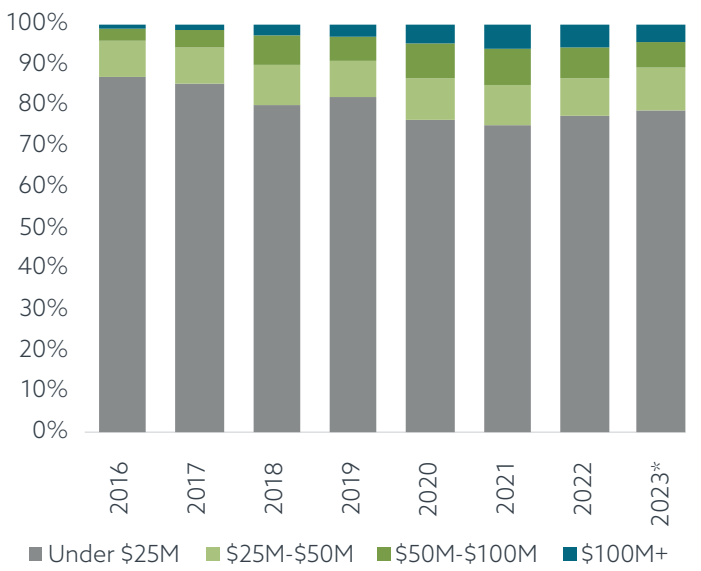
Source: PitchBook | Geography: US
*As of June 30, 2023

Share of life sciences VC deal value by size range



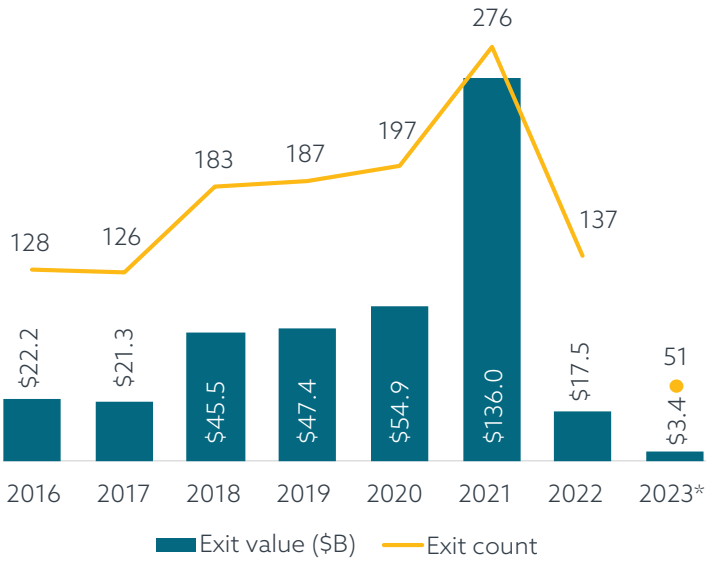
Source: PitchBook | Geography: US
*As of June 30, 2023

Share of life sciences VC deal count by size range



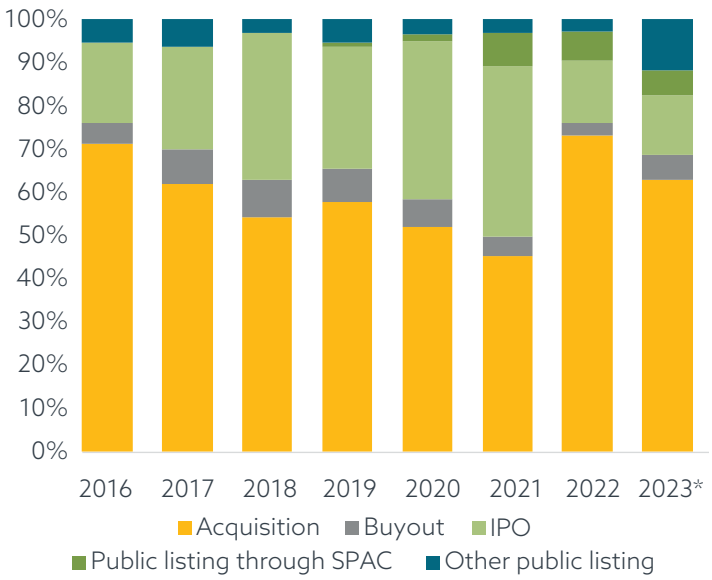
Source: PitchBook | Geography: US
*As of June 30, 2023

Life sciences VC exit activity



Source: PitchBook | Geography: US
*As of June 30, 2023

Share of life sciences VC exit count by type



Source: PitchBook | Geography: US
*As of June 30, 2023

Roundtable

Panel

Contributors

Dan Wee
Chief Of Staff, Enveda Biosciences

Systematically translating molecules found in medicinal plants into new drugs for challenging diseases. A platform harnessing nature's complexity with the help of cutting-edge advancements in knowledge graphs, machine learning, and metabolomics.

Shaq Vayda
Principal, Lux Capital

VC that is active in this space and oversees investments in tech bio companies such as Enveda, Atavistik, Creyon, Anagenex, Gandeeva, and LabGenius each of which combine artificial intelligence, robotic automation and high throughput biology to discover next-generation therapeutics.

Anna Huyghues-Despointes
Chief Corporate Development Officer, Owkin

Owkin is an AI Biotech that uses AI to unlock complex biology to find the right treatment for every patient. We integrate the best of human and artificial intelligence to deliver better drugs and diagnostics at scale. The company was founded in 2016 and has now raised over $300 million and become a unicorn through investments from leading biopharma companies (Sanofi and BMS) and venture funds (Fidelity, GV and BPI, among others).

Facilitators

Thora Johnson
Partner and Co-Chair of Life Sciences Group, Orrick

Neel Lilani
Global Head of Tech Clients, Orrick

Stephen Thau
Partner and Co-Chair of Life Sciences Group, Orrick

INTRODUCTION

Artificial Intelligence (AI) and related technologies, including how new drug therapies are created and brought to market, are increasingly prevalent. These new therapies will transform many aspects of patient care, outcomes, and speed to market, as well as administrative processes within provider, payer, and pharmaceutical organizations.

Neel Lilani: Welcome everyone and thank you for joining us today. To kick things off, I'd be interested to hear everyone's thoughts on specific areas of drug discovery that have seen the most impactful advancements through the application of (AI).

Shaq Vayda: Whenever we talk about AI and drug discovery, everyone likes to make it seem like it's this magic single tool that will fix all our problems. But unfortunately, the reality is much more complex. Every biotech company incorporating AI is probably thinking about it in a slightly nuanced and different way. On one end you have target identification: where in the body do you go after? And what is the "lock," like with the analogy of locks and keys. There are all these locks across the human body. So, how do I identify a novel, unique lock that we didn't even know existed? I can also be helpful in coming up with the design of the molecular thing that you want to hit that lock with: the key. The takeaway is that there are a lot of different ways AI is being used. It's always important to contextualize exactly what the job is that needs to be done, and how can AI help achieve that.

Anna Huyghues-Despointes: In addition to pre-clinical use cases, I would say that AI is used in clinical development. You have more methodologies to design a clinical trial, to stratify patients to improve the efficacy analysis to create external control arms as well.

Daniel Wee: From our side, a huge advancement is taking a lot of the high throughput systems and being able to and generate high-quality data at scale that can be used in an AI model. We have a lot of different methodologies to scale across the various forms of how drugs are discovered, whether it be a phenotypic assay that you're doing machine learning on, or, as Shaq mentioned, this lock and key approach to targeting a specific protein. We're taking a lot of these fundamental biological processes in different modalities of doing research, scaling them up and translating that data very quickly across many different areas. It's still early days and the proof is ultimately what the drugs come out, but it's very exciting time for the industry.

Neel: Are there any AI algorithms or models that any of you are particularly excited about?

Anna: A few years ago, there was the use of knowledge graphs across pretty much anything that was drug discovery. Every company had a proprietary graph and/or open-source software. Now, we are in the next phase of this with the large language models (LLMs). There are many models, such as Chat GPT, some cater to biology, others have agnostic use cases, most of them focus on one single data type. (structured data, protein data…) At Owkin, we have built proprietary AI algorithms and the key for us is being able to feed different data modalities into our AI engines: not only relying on one single data type. That's what's challenging in medical research. We are not going after a single use case and a single data type, therefore we cannot use off-the-shelf LLMs. We across medical research, from diagnosis, to target discovery, coumpound discovery, drug positioning and combinations, and clinical development.

Stephen Thau: Anna, how much of the potential success of AI for drug discovery is keyed to proprietary data sets versus public data sets? Which is more important, proprietary datasets or proprietary algorithms?

Anna: Let's start by answering the data aspect of your question. The companies that are going to be successful in the long term are the ones that will continue to evolve the data on which they work. Current cohorts need to be augmented by adding novel data modalities such as spatial omics. Novel data will unlock a novel biological signal. You must get the latest technology to generate it. Continue to invest in patient cohorts also means to understand the efficacy of the latest treatments. Once you make a druggable discovery you need to elaborate the development plan, the trials, the line of treatment you are envisioning and competitive products. And novel treatments move through the treatment lines (lately anti-PD1s having been moving forward. And now on the algorithm side, it's true that AI is being democratized, and large language models are enabling discovery and applications. But again, the complexity in biology is in the link of different types of data sets together, and today, there is not one model that will be able to integrate clinical data, DNA, proteins, cells, tissue images, chemistry… So public AI can be used for specific use cases but does not generalize across translational research and drug development.

Thora Johnson: Anna, on that note, can I ask you a question about the federated learning technology at Owkin?

Anna: It's an infrastructure which is the same across all the hospitals and the partners that we work with. That infrastructure, called Substra, is hosted by the Linux Foundation, and is open source. It's the backbone. It enables you to do two things. Firstly, you can access the data remotely, meaning that you don't have to extract the data out of the partner infrastructure. You send the models to train locally, and then you only get back the model, not the data. Secondly, it enables the federated learning, which is this remote training that is done across multiple nodes at the same time. The federated learning is the action of training a machine learning model across multiple data sets at the same time, protecting the data and the sharing of the data. Substra is privacy preserving and compliant with the latest patient data regulations in Europe and the US. Everything that we do at Owkin is done that way. And Substra is agnostic, it can deployed on top of any infrastructure, data can be stored on premise or in the cloud.

Neel: And the data limitation is that these are not publicly accessible data sources, and they're sort of siloed? And the similar models, such as OpenAI, cannot access them to evaluate them?

Shaq: Exactly. Every research institution and company obviously hold information proprietary. There isn't a treasure trail of the Internet text that exists for protein database, and that obviously is manifested in a really compressed model. So now we wonder if there could be an OpenAI for biology.

Daniel: At Enveda, we just released this preprint called MS2Mol, which is based on a transformers model, which is what ChatGPT is build on, and happens to work particularly well for mass spectra data sets. This is because the context of your peaks in it is dependent on the actual other fragments of the molecule, much in the same way that a sentence takes on different meanings based on the order and the context of the words in it. This has worked out phenomenally well for us, being able to predict chemical structures, which is a huge blocker for looking into unknown libraries and generating hits from natural products which have been successful in developing new drugs. Then, going into the data component, mass spectra has been around for decades. We can leverage a lot of what's publicly available, but it's always a marriage: the biology and the data quality cannot be separated when you're looking at this space.

Shaq: One really important point that Dan touched on is language. We take it for granted, but there is an inherent prediction: the word I just said previously has kind of finitely made the universe of the next possible word smaller. That's how this fundamentally works. They predict the next word. Even with something like amino acids, we don't actually know what the next logical amino acid should be given the immune essence before it, so we're actually trying to apply what has worked in one domain, hoping we can generalize biology. But maybe there's another pre-training process that we haven't discovered yet.

Thora: I was going to ask about hallucinations and the use of generative AI. How do you control for it, or what are you thinking about in terms of hallucinations?

Daniel: To a certain extent, all life science companies are trying to understand biology a little more. You actually want a generative AI model that is able to explore and not be confined by very strict lanes, because we don't fully understand the rules that are governing the biology in itself, and part of that is hallucination. False information is part of that journey of whittling down. What are your parameters? What are your lanes that you want to look at? It also depends on what stage of the process you're in.

Shaq: At the end of the day, I think it's important to contextualize that all of these tasks are in service of building a drug. You might generate a molecule, but you still need to test it and see how the outcomes measure up to the predictions. If they don't, then you go back and retrain the model. I don't fear the hallucination, at least in that in the areas we're talking about.

Anna: I agree, we always validate the discovery that our AI engines make. We validate across cohorts and then we validate with experiments. Everything has to be validated from a methodology perspective and from a biological perspective.

Stephen: Do you see a time on the horizon where AI eliminates the need for some wet biology testing?

Shaq: There's definitely an excitement. Who knows if in five or ten years from now we could be there, but I'm skeptical. I think we're still so early on in our understanding of AI's potential that I don't really feel that the black box mindset is a realistic one.

Daniel: The true value of AI is that it can guide scientists into understanding what experiments to run, and what prioritizations to make. Essentially, it's shortening the iteration cycle rather than eliminating it completely.

Anna: I agree. Relying 100% on AI is not the point. AI is helping to get to humans faster. There are novel technologies that are mimicking organs to get more information on how treatment modulate specific targets and the tumor environment. These organoids are generating large amounts of data that can be analyzed with AI as well. The more complex the experimental data becomes the more there is a need for AI.

Neel: Dan, how you are thinking about the ethical implications of leveraging AI in drug discovery?

Dan: Our approach is unique, and the way that we apply AI is early in the drug development/discovery cycle. Our AI platform predicts the chemical structures that comes out of our hits. So, when it comes to predicting efficacy and safety, we rely on the fact that natural products having a tremendous success rate. However, this does bring up something that I think is commonly applied towards other types of companies who are looking at natural sources, whether it be where their libraries generate from, or even, say, variance in humans genotypes and looking at potential targets from that perspective. For us, a large amount of consideration is directed towards ethically sourcing our library. There is a multilateral treaty known as the Convention on Biological Diversity that has established an agreement called the Nagoya Protocol [on Access to Genetic Resources and the Fair and Equitable Sharing of Benefits Arising from their Utilization to the Convention on Biological Diversity], which offers a benefit share for products that were derived from either the cultural use or the natural use of the countries of origin. Many countries have ratified this convention. However, we uncovering the vast unknown chemical space in natural products and are seeing that there are a lot of overlaps in chemistry and their uses across countries and culture. We are focused using what we learn to improve the effectiveness of the Nagoya Protocol and on helping countries realize the value in their biological diversity.

Neel: Shaq, from an investor perspective, what are the primary ethical considerations of AI that influence your investment decisions?

Shaq: AI alignment and AI safety are very topical right now. We're seeing it converge from investors saying there's an expectation that this should be part of the diligence process, and you've got the public figures who are also thinking about it from regulatory standpoint, making sure that it's being incorporated. There is sort of a fear, though. The open-source community is building and experimenting, continuing to publish papers, and the onus is on the actual companies themselves. They incorporate these models to understand the data sets. Are they representative of larger patient samples? By understanding a lot of that, quantifying it, we're building more tools that can help us with those tasks.

Thora: Have you seen companies leverage or make reference to the NIST protocol for AI, and trustworthiness of AI? Do you think it has any relevance in life sciences or for companies to even be thinking about it?

Shaq: It feels like everyone is sort of there. A lot of research organizations are popping up with claims about what AI alignment looks like and how you build. I think a lot of the top-down regulations will probably drive some of the focused thinking around this as well.

Anna: I think authorities have gotten to the point where it is well understood that patient health care data is very valuable, and it needs to be regulated and accessed in a privacy preserving way. Europe is more advanced on this topic than the U.S., but the U.S. still has some regulation in place. Other countries such as China and India have large patient population and are in process of centralizing health care data to use in research and commercial settings. We except regulations to synchronize across countries.

Thora: You're mapping your use of AI, and then you're managing it and monitoring it. It's very abstract.

Shaq: This idea of bio security is popping up as an investment theme, and really the same models that we're using to optimize properties can also be used to optimize properties that aren't necessarily beneficial for a therapeutic. So, the same tool that can be used to design a therapeutic is also a pathogen or can be used in some other kind of negative context. There's a lot of scrutiny right now around how we make sure that all actors are intending to use these things for good.

Neel: Shaq, from an investor lens, what are the major criteria investors are prioritizing when thinking about where to deploy capital in the AI ecosystem?

Shaq: It's an interesting question, especially given the macro trends that are happening. We think about, how ultimately the mandate is to invest in innovation, and we fundamentally believe this innovation will compound and be a massive creator of value in the long term. I think the life sciences sphere is one of the most interesting places where AI will ultimately be used.
I think we're going to see increasingly more and more investors spending time in the physical world, especially now that some of these systems have found their way to help provide efficiencies and cost savings for workflows that historically would not have been able to be optimized. I become increasingly excited about the stuff we're already investing in. I believe that a lot of the companies that are building, whether it's therapeutics or biomarker discovery, or things that ultimately require hard problems, those are where more and more investors are going to spend time.

Neel: Dan and Anna, both from the company side: how have you seen your fundraising process evolve?

Anna: We have a different equity story than a lot of companies, because we started as a pure machine learning company. Google Ventures was the first investor, and then we started servicing pharma with drug discovery and from development programs. There, we attracted large health care investors. One we had a stable alliance business, we decided to start developing our own pipeline. We build a successful business with the pharma partnerships and pharma strategic alliances, we are now cash flow positive on this business. As we launch our proprietary pipeline, we will involve a different class of investors focused on therapeutics vs. AI or diagnostic. For us, it is a question of how we think about the next step for each side of the business: pharma services, therapeutics, diagnostics. It's not one size fits all.

Daniel: As Shaq knows, we never stop talking to investors. We continue to keep our ears on the ground, getting a pulse on what their thoughts are. It's hugely insightful. At Enveda, we were founded right before the pandemic and throughout and beyond the pandemic I think a lot of the macros led to a huge roller coaster ride in terms of what that broader fundraising environment has been. As we got started, a lot of our pitch was focused on our platform, because that was all we had. Then, as soon as we got assets, the diligence became more mature, and now we are at a stage where we're entering the clinic and investors are looking at it from a completely different lens, evaluating the potential of our portfolio. It's ultimately maintaining that line of sight and having your platform and team deliver better drugs faster and cheaper. I think that's always held true and has worked towards our benefit and our success in fundraising as the market has been volatile.

Stephen: What are some things that you'd anticipate happening in the next three to five years in AI?

Shaq: There's a recently published paper about the largest ever protein train model. It outperformed the state-of-the-art protein model on 13 out of 15 tasks. Every day, we're seeing something brand new pop up. I think we're about to see a real revolution in the Golden Age medicine, as the New York Times call it. That's all to say that the confluence of data, the confluence of automation, and the confluence of generally better tools will allow us to integrate biology to a degree that never has happened before. It's a really great time to be in this industry.

Daniel: It is very early days in AI and drug discovery. It's only a matter of time until we see drugs that were discovered using AI performing well in patients, but also seeing that AI approaches to drug discovery outperform traditional approaches on cost, time, and failure rates. We are also seeing tremendous advances in Foundation models and their applications to chemistry and biology. This will lead to more and more big pre-training approaches in vision and language as applied to drug discovery. Ultimately, there are so many verticals in the drug discovery and development that can benefit from AI and what you will see are more companies developing AI tools to address more specific parts process that will not only accelerate the time to market, but also allow companies, specifically startups, to do more with less. What I hope and predict is that this will lead to an unprecedented number of drugs being developed for the vast number of diseases which have no treatment at all.

Anna: Anna: I anticipate AI diagnostic tools to become widely available and used in clinical settings and in testing labs. I also see automation of wet lab activities to generate experimental data at scale that will fuel target and drug discovery. Finally, novel sequencing and imaging technologies will enable to unlock novel signal in patient data.

CREATORS. VISIONARIES. UNDERDOGS. THE RISK TAKERS.

We help you disrupt. We help you grow.
We protect you. We help you win.

Top 3 for Innovation 7 years in a Row 2016-2022
— Financial Times

#1 in Reinventing Legal Practice 2022
— Financial Times

#1 Most Active Venture Capital Law Firm in Europe – 29 quarters in a row!
— PitchBook Q1 2023 League Table Rankings

#5 Globally – Top 5 Globally for 23 quarters in a row!
— PitchBook Q1 2023 League Table Rankings

CONTACT US

Carsten Bernauer
Partner
cbernauer@orrick.com

Tony Chan
Partner
tychan@orrick.com

Ed Dyson
Partner
edyson@orrick.com

Craig Falls
Partner
cfalls@orrick.com

Gregg Griner
Partner
ggriner@orrick.com

Blake Ilstrup
Partner
bilstrup@orrick.com

Scott Iyama
Partner
siyama@orrick.com

Thora Johnson
Partner
thora.johnson@orrick.com

Neel Lilani
Global Head, Tech Clients
nlilani@orrick.com

Ed Lukins
Partner
elukins@orrick.com

Mike O'Donnell
Partner
mike.odonnell@orrick.com

David Schulman
Partner
dschulman@orrick.com

Shana Solomon
Partner
shana.solomon@orrick.com

Gargi Talukder
Partner
gtalukder@orrick.com

Stephen Thau
Partner
sthau@orrick.com

Albert Vanderlaan
Partner
avanderlaan@orrick.com