PitchBook

PitchBook Data, Inc.

Nizar Tarhuni Executive Vice President of Research and Market Intelligence

Paul Condra Global Head of Private Markets Research

James Ulan Director of Emerging Technology Research

Institutional Research Group

Analysis

Ben Riccio
Associate Research Analyst
ben.riccio@pitchbook.com

Data

Harrison Waldock
Data Analyst

pbinstitutionalresearch@pitchbook.com

Publishing

Designed by **Drew Sanders**

Published on November 12, 2025

Contents

EMERGING TECH RESEARCH

AI in Drug Development

Tracking AI's growing influence in biopharma

PitchBook is a Morningstar company providing the most comprehensive, most accurate, and hard-to-find data for professionals doing business in the private markets.

Key takeaways

- **AI has proliferated across the biopharma ecosystem:** Startups are deploying AI at every step of the drug development process: developing their own pipelines of AI-discovered assets, inking massive partnership contracts for discovery services, offering AI-enabled platforms and services, and deploying AI in clinical trials. While approaches vary, each is aimed at addressing the fundamental challenges of drug development: high failure rates, rising costs, and lengthy development timelines.

- **Biology shifts to foundation models:** The introduction of foundation models in biology has already transformed protein-structure prediction and bioinformatics. Frontier applications focus on "virtual cells": comprehensive simulations of cellular behavior capable of modeling therapeutic response and toxicology before clinical-trial validation.

- **VC activity picks up:** Over the past 12 months, VC investment in AI drug development has totaled $3.2 billion across 135 deals. AI-enabled pharma tools & services constitute 62.5% of the total deal count, while AI-native drug developers have attracted the larger share of the total capital, at 70.1%.

- **AI-native biotechs have a nearly 100% valuation premium:** In 2024, AI-native biotech companies raised money at a median valuation of $78 million, nearly twice the broader biopharma median. The valuation gap highlights both the frenzied AI investment environment and investor conviction in the ability of AI integration to drive greater R&D efficiency and improve clinical success.

- **AI pharma tools & services see dealmaking flurry:** AI pharma tools & services have seen a significant uptick in VC activity over the past four quarters, with deal counts up 51.8% on a TTM basis. Startups in the space offer a faster path to profitability and more asset-light business models, desirable characteristics in the current biopharma slowdown.

- **Data limitations present a barrier to scaling:** With early biological foundation models showing evidence of scaling behavior, continued progress hinges on overcoming data siloes and quality limitations. A competitive advantage in new model development will come from the ability to generate or acquire large-scale, high-quality biological datasets.

Overview

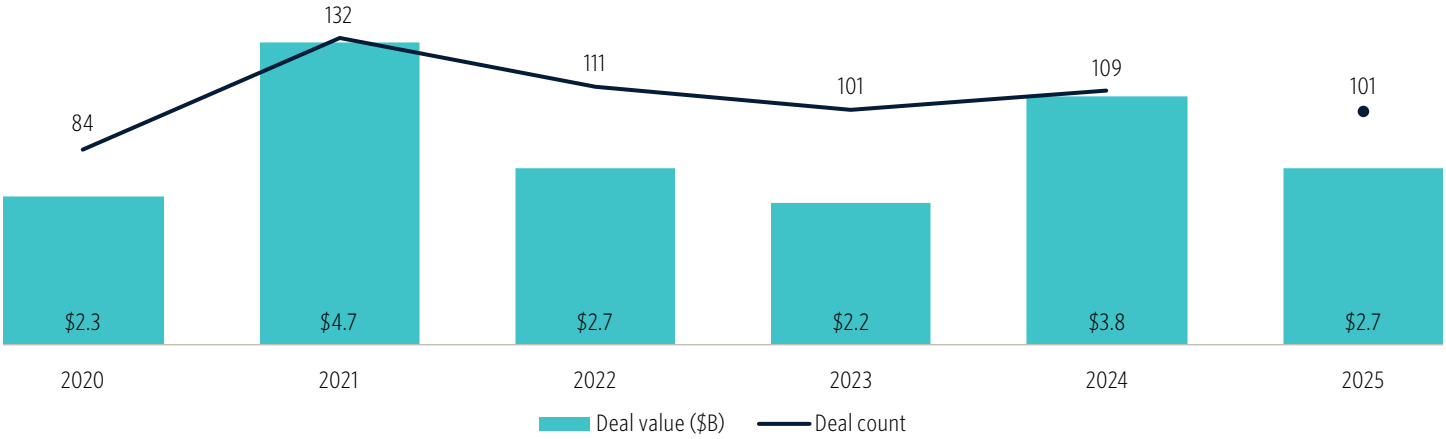
While AI is reshaping nearly every industry, its influence may prove most transformative in biopharma, where new technologies promise to address the high failure rates, lengthy development timelines, and escalating costs that characterize the industry. While these goals remain lofty, significant progress has been made in recent years, with frenzied investment in AI and a massive computing build-out fueling advancements in research productivity, drug discovery, and clinical-trial efficiency.

The current iteration of AI in drug development is driven by the application of foundation models in biology. These models, trained on vast biological datasets, have built upon the capabilities of earlier machine learning applications in chemistry that were developed over a decade ago and commercialized by startups such as Atomwise and Insilico Medicine. These smaller models used task-specific architectures and were trained on labeled datasets, making them valuable tools for target identification and screening, but they had limited capabilities outside these tasks. Newer models are trained on large unlabeled datasets, enabling generalization across tasks and an improved ability to make predictions and design de novo molecules. These tools have led to several breakthroughs and spurred a new class of AI-native biotech startups, which are among the most well-funded startups in the biopharma sector. The next frontier of foundation models lies in integrating data across biology to produce comprehensive simulations of biological systems, further shifting drug discovery in silico.

AI technologies have emerged across the biopharma ecosystem: Startups are developing their own pipelines of AI-discovered assets, inking massive partnership contracts for discovery services, offering AI-enabled platforms and services, and deploying AI in clinical trials. While approaches vary, all these startups are looking to address the structural issues within drug development. Bringing a drug to market can take over 10 years and $1 billion, and with a clinical success rate of around 10%, failure is common.[1] These challenges have contributed to one of the most drastic funding downturns for the sector in over a decade, as a general risk-off sentiment among VC investors has pushed capital away from early-stage biopharma. A record-low 14.5% of all biopharma deals YTD have gone to pre-seed/seed startups, compared with the previous five-year average of 17.3%. Investors are instead prioritizing clinical-stage assets or moving away from the sector entirely to pursue faster growth opportunities in other tech verticals. AI offers a potential solution, de-risking drug discovery and reducing its cost basis. The technology's ability to transform research & development (R&D) and increase efficiency in a traditionally inefficient sector may be crucial to fostering the return of early-stage deal activity and reigniting VC interest in the sector.

1: "Why 90% of Clinical Drug Development Fails and How to Improve It?" National Library of Medicine, Acta Pharmaceutica Sinica B, Duxin Sun, et al., February 11, 2022.

AI drug development VC deal activity



Source: PitchBook • Geography: Global • As of September 30, 2025

Top VC investors in AI drug development companies since 2020

Investor	Investor type	Total deal count
Lux Capital	VC	19
5Y Capital	VC	19
Gaingels	VC	17
Alumni Ventures	VC	15
Andreessen Horowitz	VC	15
OrbiMed	Growth/expansion	15
Khosla Ventures	VC	14
Catalio Capital Management	VC	13
HongShan Capital Group	VC	13
MSD Global Health Innovation Fund	Corporate VC	13

Source: PitchBook • Geography: Global • As of September 30, 2025

Technologies and recent developments

Protein-structure prediction and de novo design

Protein structure defines biological function, and accurate prediction of three-dimensional conformations enables researchers to better understand disease states and develop more targeted drugs. AI has revolutionized protein-structure prediction, overcoming a major experimental bottleneck in biology. DeepMind's AlphaFold 2 achieved a breakthrough at the 2020 protein-structure competition CASP14, achieving near-experimental accuracy. In 2021, DeepMind open-sourced the model and released structure predictions for over 200 million proteins. As traditional experimental methods for structure determination can take months or years and cannot be applied to all proteins, the model and the following dataset greatly expanded the range of druggable targets and accelerated target identification.

Subsequent iterations of protein models have expanded their functionality. Meta's ESMFold achieved faster inference, enabling less computationally intensive analysis of larger datasets, while AlphaFold 3 added modeling of protein-ligand and protein-nucleic-acid interactions, furthering its application in drug design.

Models have also moved beyond structure prediction to de novo design. RFdiffusion, developed by the Baker Lab, applies diffusion-based generation—like those used in image and video models—to create entirely new proteins with desired structures and functions. Results have included AI-generated binding proteins, complex enzymes, and, most recently, ion channels.[2, 3] Other breakthroughs include Germinal, an open-source model from the Arc Institute and Stanford University capable of generating antibodies with high binding affinity to a desired target, and OpenCRISPR-1, a novel Cas9-like enzyme with improved gene-editing functionality developed with proprietary models by the startup Profluent.

Genomic models

AI has also become increasingly valuable in genomic analysis. Early tools such as DeepVariant and DeepConsensus improved the accuracy of variant detection and the quality of sequencing data; however, these models lacked the scale to generate insights in a wider genomic context. The emergence of transformer-based models expanded this capability, allowing models trained on multispecies genomic datasets to capture evolutionary conservation patterns and functional elements. Models such as DNABERT demonstrated that foundation models could be used to identify regulatory elements and predict variant effects through fine-tuning.

2: "Computational Design of Serine Hydrolases," Science, Anna Lauko, et al., February 13, 2025.
3: "Bottom-Up Design of Ca^{2+} Channels From Defined Selectivity Filter Geometry," Nature, Yulai Liu, et al., October 22, 2025.

Importantly, these models appear to follow similar scaling laws to those observed in language and vision models, where performance improves with dataset and model size. More recent models have dramatically expanded their scale. Evo 2, developed by the Arc Institute and NVIDIA, was trained on 9.3 trillion base pairs from over 100,000 species and has a 1-million-nucleotide context window, enabling the analysis of long-range genomic interactions.[4] The model can identify pathogenic mutations, annotate genomes across species, and generate sequences with defined chromatin accessibility. Researchers later applied the model to design a functional bacteriophage genome, illustrating both the potential of AI in synthetic biology and its emerging biosecurity threats.[5] Taken together, these advancements are transforming genomics into a software-like industry, where innovation is no longer constrained by wet-lab throughput. Still, questions of algorithmic bias, data privacy, and equitable access remain pressing as these tools move closer to clinical deployment.

Cell foundation models

The next frontier for AI models in biology is the development of comprehensive simulations, or "virtual cells," capable of modeling intracellular interactions and cellular response to perturbation—enabling in silico therapeutic validation. Foundation models trained on single-cell RNA sequencing data have emerged as an initial approach. Models such as scGPT, Geneformer, and Cell2Sentence can classify cell types based on a given transcriptome or predict gene-expression changes induced by various perturbations. However, studies have found that these models underperform compared with simple baseline methods.[6, 7] Achieving truly predictive "virtual cells" that can model efficacy and toxicity across diverse patient populations will likely require multimodal integration across transcriptomic, proteomic, and spatial datasets—areas currently limited by siloed and noisy data.

AI-native biotech

AI drug discovery

As the capabilities of AI models continue to increase, a new class of AI-native biotech companies (biotechs) has emerged.[8] These startups have rapidly become some of the most valuable in the sector, with Xaira Therapeutics, Isomorphic Labs, and Generate:Biomedicines all reaching unicorn status within their first two institutional rounds. By building and leveraging internally developed models, these startups promise to improve clinical success rates, reach a wider range of druggable targets, and operate more efficiently than traditional drug developers.

4: "Genome Modeling and Design Across All Domains of Life With Evo 2," bioRxiv, Garyk Brixi, et al., February 21, 2025.
5: "Generative Design of Novel Bacteriophages With Genome Language Models," bioRxiv, Samuel H. King, et al., September 17, 2025.
6: "Deep-Learning-Based Gene Perturbation Effect Prediction Does Not Yet Outperform Simple Linear Baselines," Nature Methods, Constantin Ahlmann-Eltze, Wolfgang Huber, and Simon Anders, August 4, 2025.
7: "Benchmarking Foundation Cell Models for Post-Perturbation RNA-Seq Prediction," BMC Genomics, Gerold Csendes, et al., April 23, 2025.
8: This analyst note defines AI-native biotechs as startups that use AI as a core component of their preclinical or clinical development process and primary source of differentiation. A full list of included companies is available on the PitchBook Platform.

Research has found an 80%-90% Phase I success rate among the early cohort of assets developed by AI-native biotechs, significantly above the industry average of 40% to 65%.[9] Phase II outcomes, however, were in line with historical averages at 40%, based on a limited sample of just 10 trials.[10] Although the dataset is small, this preliminary research suggests a doubling of the overall probability of success, from approximately 8% to 16%, when historical phase-by-phase benchmarks are modified. Additionally, by lowering the cost and time frame of discovery, AI-native biotech companies can produce more shots on goal without escalating costs—a structural de-risking, even if future efficacy outcomes converge with traditional success rates. Ongoing monitoring of these preliminary figures will be critical in determining the impact of AI in later phases, where success is more governed by complex biology.

In addition to offering comparatively more efficient internal development, AI-native biotech startups, like other platform companies, are closing large deals to collaborate with established drug developers. With several deals over $1 billion, when including milestone payments, these partnerships (listed in the table "Select partnership deals for AI-native biotech startups") underscore Big Pharma's belief in the technology and provide early validation before internally developed assets reach the clinic.

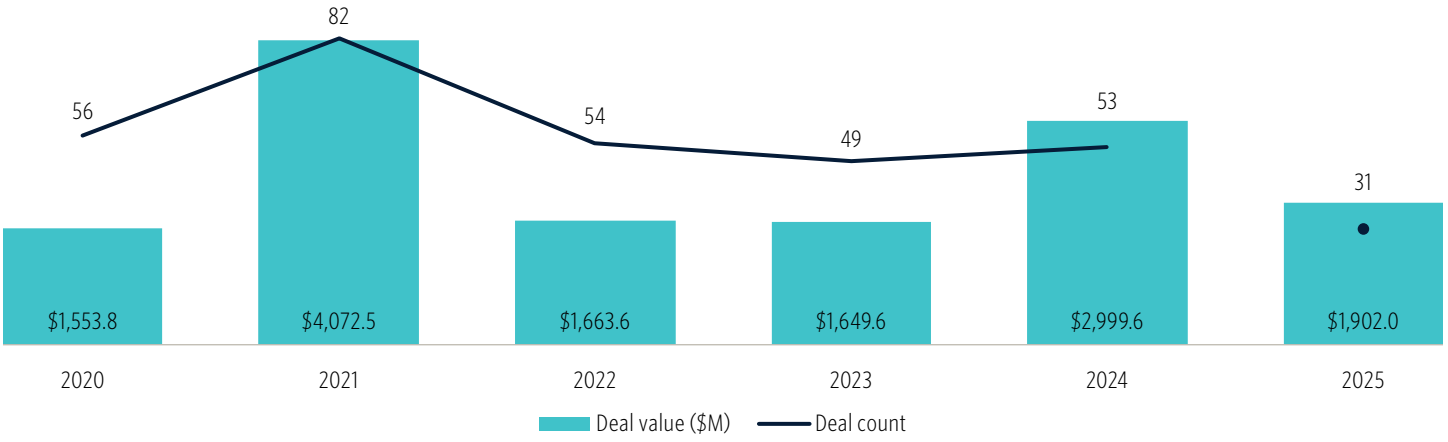
VC funding for AI-native biotechs

Alongside the broader biopharma sector, VC funding for AI-native biotech startups peaked in 2021, with $4.1 billion invested across 82 deals. Through the first three quarters of 2025, the space has attracted 31 deals totaling $1.9 billion. While AI's share of the total biopharma deal count has remained around the five-year annual average of 4.1%, its share of the total capital has increased over the past two years, landing at 9% in 2024 and 8.4% YTD. This reflects the increased occurrence of large deals in the space, with Xaira Therapeutics having raised over $1 billion since 2024 and Isomorphic Labs attracting $579.1 million earlier this year, as well as the ability of these startups to raise larger rounds than traditional biotechs as AI continues to capture an outsized share of VC attention. Still, AI remains a minor area of investment in biopharma compared with the broader venture ecosystem, where AI startups have captured 64.3% of total venture dollars and 37.5% of deals YTD. The trend reflects the slower pace of AI integration in biology and a broader investor preference for more market-ready applications in other tech verticals.

A premium emerges when comparing median pre-money valuations of AI-native biotechs with those of the larger biopharma sector. The median pre-money valuation for AI-native biotechs was nearly twice that of the broader industry in 2024 and remains elevated at 1.4x YTD. Median deal sizes for AI-native biotechs are also larger: 1.9x the industry median in 2024 and 1.6x in 2025. This indicates that AI-native startups are raising larger rounds at heightened valuations compared with traditional biopharma companies—unsurprising in the current AI investment environment.

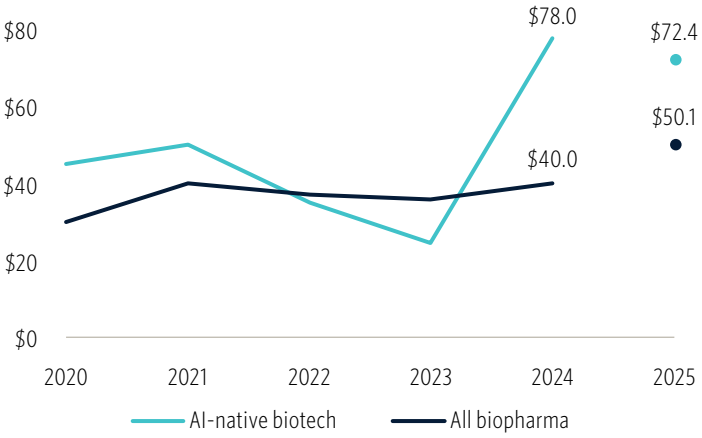
9: "How Successful Are AI-Discovered Drugs in Clinical Trials? A First Analysis and Emerging Lessons," ScienceDirect, Drug Discovery Today, Madura K.P. Jayatunga, et al., June 2024.
10: Ibid.

AI-native biotech VC deal activity



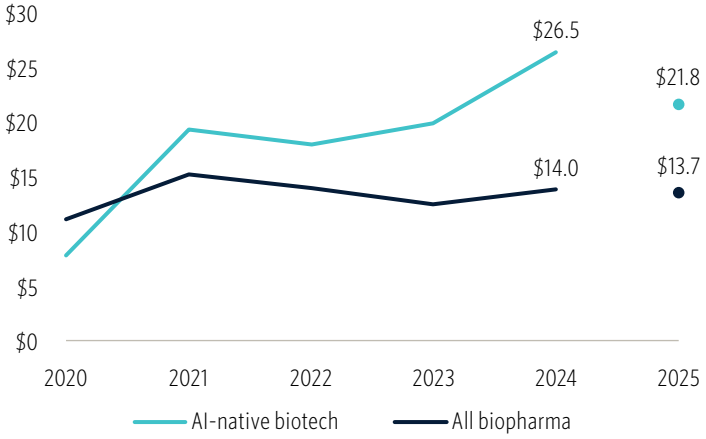
Source: PitchBook • Geography: Global • As of September 30, 2025

Median biopharma VC pre-money valuation ($M) by company type



Source: PitchBook • Geography: Global • As of September 30, 2025

Median biopharma VC deal value ($M) by company type



Source: PitchBook • Geography: Global • As of September 30, 2025

Top AI-native biotech startups by total VC raised

Company	VC raised to date ($M)	Highest clinical trial phase	Therapeutic area(s)	HQ location	IPO probability	M&A probability
Xaira Therapeutics	$1,000.0	Preclinical	N/A	South San Francisco, US	91%	7%
Generate:Biomedicines	$693.0	Phase 1	Immunology & infectious disease	Somerville, US	95%	3%
insitro	$643.2	Preclinical	Neurology, metabolic	South San Francisco, US	81%	11%
Formation Bio	$618.8	Phase 3 (out-licensed)	Immunology	New York, US	87%	11%
Isomorphic Labs	$579.1	Preclinical	N/A	London, UK	N/A	N/A
Insilico Medicine	$549.3	Phase 2	Systemic & structural disorder	Boston, US	88%	10%
Valo Health	$530.0	Phase 2	Ophthalmology	Lexington, US	97%	1%
Enveda	$530.0	Phase 1/2	Immunology	Boulder, US	95%	3%
Pathos	$467.0	Phase 2	Oncology	Chicago, US	85%	9%
Genesis Therapeutics	$327.4	Preclinical	Oncology	Burlingame, US	89%	9%

Source: PitchBook • Geography: Global • As of September 30, 2025
Note: Probability data is based on PitchBook VC Exit Predictor methodology.

Select partnership deals for AI-native biotech startups

Company	Pharma partner(s)	Deal year(s)	Total potential deal value(s)	Partnership focus area
Valo Health	Novo Nordisk	2023-2025	$4.6 billion	Partnership expanded in 2025 for discovery of up to 20 programs across cardiometabolic conditions
insitro	Bristol Myers Squibb	2025	$2.0 billion	Partnership, which began in 2020, for discovery of small-molecule treatments for ALS
Generate:Biomedicines	Amgen, Novartis	2022, 2024	$1.9 billion, $1.0 billion	Two deals: Discovery of up to 10 new drugs with Amgen, discovery of protein therapeutics with Novartis
Isomorphic Labs	Eli Lilly, Novartis	2024	$1.7 billion, $1.2 billion	Two deals, both for discovery of small-molecule therapeutics
Superluminal Medicines	Eli Lilly	2025	$1.3 billion	Discovery of small-molecule drugs for cardiometabolic disorders
Insilico Medicine	Sanofi	2022	$1.2 billion	Access to AI platform for six drug candidates
Atomwise	Sanofi	2022	$1.2 billion	Access to AI platform for up to five small-molecule candidates
Creyon Bio	Eli Lilly	2025	$1.0 billion	Discovery of oligonucleotide therapeutics
Algen Bio	AstraZeneca	2025	$555.0 million	Access to AI platform for discovery of targets in immunology
Manifold Bio	Roche	2025	$2.0 billion	Discovery of drug delivery molecules for neurology
Nabla Bio	Takeda	2025	$1.0 billion	Discovery of protein therapeutics

Source: PitchBook • Geography: Global • As of September 30, 2025

AI pharma tools & services

AI lab assistants

Modeling vertical AI applications in other tech sectors, biopharma startups are deploying models trained or fine-tuned on scientific literature and biomedical data that is often siloed and unstructured to offer laboratory-specific capabilities. These specialized AI lab assistants can improve experimental workflows by generating protocols, suggesting reagents, and informing target selection. The result is a faster, more informed R&D cycle, where researchers can incorporate a significantly wider breadth of scientific evidence when designing experiments.

These tools offer a more asset-light business model than typically seen in biopharma or other pharma-tools startups. However, like other vertical AI applications, they face inherent competition from the growing scope of leading foundation model developers. Anthropic recently announced Claude for Life Sciences, which offers connections to several popular platforms such as Benchling and PubMed, as well as new skills in single-cell RNA sequence data analysis. While these capabilities are still limited, the moves signal that AI-enabled life science platforms must be highly specialized and offer unique features to compete with horizontal AI models.

BenchSci is one of the leading companies in this space, raising $166.7 million in venture capital to date and establishing partnerships with Thermo Fisher Scientific and Sanofi. The startup uses text and vision models to inform reagent selection, map disease biology, and improve preclinical discovery. Causaly, another major startup in the space, focuses on biomarker identification and target prioritization through scientific literature review.

Discovery services and model developers

Startups are also pursuing an AI-enabled services approach, providing access to proprietary AI platforms or establishing drug development collaborations. Unlike with the previously discussed partnerships, these startups do not develop their own assets internally and instead operate solely as service providers. With data siloes and limited high-quality biological datasets, value in this segment is derived from generating valuable training data and providing access to AI platforms.

Lila Sciences is one of the largest and most ambitious startups pursuing this business model. After emerging out of Flagship Pioneering earlier this year, the startup has raised $435 million in VC funding, most recently at a $1.3 billion post-money valuation. The company is pursuing "scientific superintelligence," developing AI systems incorporated with lab robotics to automate the scientific process. By addressing data bottlenecks through an automated AI system, the company enables scalable model improvement and accelerates the pace of biological discovery. Lila has stated that it is not developing its own therapeutics but instead is positioning itself as a service provider.[11] Other notable startups include Profluent, which leverages protein-design models for discovery or out-licensing partnerships, and Chai Discovery, which offers a generative antibody design platform.

11: "Flagship Startup Raises $200M in Pursuit of 'Scientific Superintelligence,'" BioPharma Dive, Gwendolyn Wu, March 12, 2025.

These startups offer shorter timelines to revenue by bypassing the drug development process but face inherent risk from the commodification of biological AI models as Big Pharma increasingly expands internal model development. Additionally, integration into established R&D workflows and validation processes may prove cumbersome, slowing adoption despite proof-of concept merit.

The full list of AI-enabled pharma tools & services startups is available on the PitchBook Platform.

AI in clinical trials

AI applications in clinical trials address a critical bottleneck in drug development. These tools can inform trial design by modeling patient populations most likely to respond to a drug, identifying key biomarkers, or selecting patients. Other AI tools can handle more administrative tasks, such as patient recruitment and proposal writing.

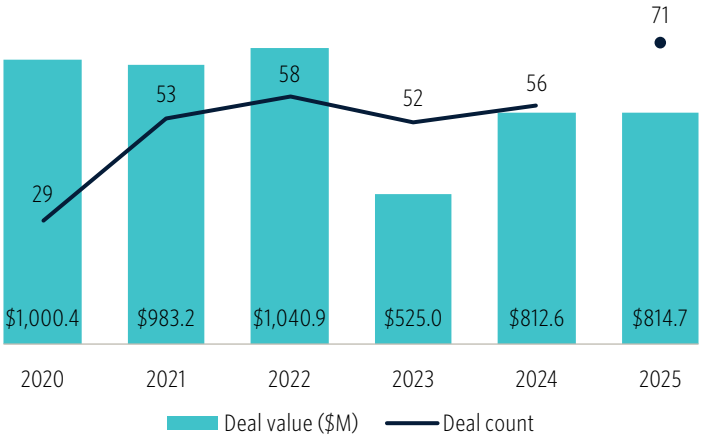
Unlearn is a leading player, raising $134.9 million in VC funding to date, most recently at a $360 million post-money valuation. Its platform uses digital twins to forecast how individuals would progress under the standard of care or placebo, reducing the size of the control arms. Other startups include PhaseV, which offers a suite of AI features for both trial design and data analysis. The company claims these tools reduce costs by 50%, decrease enrollment size and trial duration by 40%, and boost the likelihood of trial success by 30%.[12]

VC funding for AI pharma tools & services

In the first three quarters of 2025, 71 deals for AI pharma tools & services startups closed on a total of over $800 million—both values on pace for record years. Strong investor interest in the emerging tech-bio space, and AI in particular, has significantly increased deal activity in the segment. Additionally, with a median deal size of $6.8 million, these platforms offer comparatively stronger capital efficiency than drug development companies—a desirable characteristic in the current investment environment.

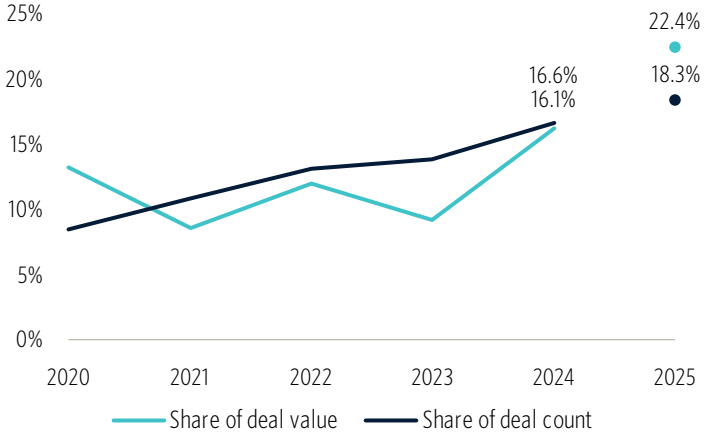
12: "PhaseV Lands $50M Series A to Supercharge AI for Clinical Development, Backed by Top VCs and Trusted by Leading Pharma Clients," PR Newswire, PhaseV, May 13, 2025.

AI pharma tools & services VC deal activity



Source: PitchBook • Geography: Global • As of September 30, 2025

AI pharma tools & services VC deal activity as a share of all pharma tools & services VC deal activity



Source: PitchBook • Geography: Global • As of September 30, 2025

Notable AI pharma tools & services companies

Company	VC raised to date ($M)	HQ location	IPO probability	M&A probability	No exit probability
Owkin	$334.1	New York, US	26%	38%	36%
Lila Sciences	$200.0	Cambridge, US	32%	66%	2%
EvolutionaryScale	$182.0	New York, US	88%	10%	2%
BenchSci	$166.7	Toronto, Canada	29%	69%	2%
Unlearn	$134.8	San Francisco, US	84%	13%	3%
DP Technology	$128.7	Beijing, China	N/A	N/A	N/A
CytoReason	$123.0	Tel Aviv, Israel	86%	10%	4%
Cradle	$102.5	Amsterdam, Netherlands	17%	39%	44%
Chai Discovery	$100.0	San Francisco, US	77%	21%	2%
BioMap	$100.0	Palo Alto, US	52%	33%	15%
Huashen Zhiyao	$84.6	Beijing, China	N/A	N/A	N/A

Source: PitchBook • Geography: Global • As of September 30, 2025
Note: Probability data is based on PitchBook VC Exit Predictor methodology.

Challenges and outlook

Challenges

Progress in AI integration in biopharma has been significant, but several challenges remain as the technology matures:

- **Data siloes:** AlphaFold 2 and similar protein models were trained on over 200,000 structures in the publicly available Protein Data Bank. This unique scenario is not reflective of the wider biological data landscape: datasets of cellular perturbations, RNA structure, toxicology, and other relevant biological

fields are highly siloed and of lower quality. As a result, many model developers are vertically integrated, utilizing their own wet labs to generate relevant training data. This imposes high costs and timelines, positioning the most well-funded startups as leaders in model innovation. However, this also presents an opportunity for startups with strong data-generation capabilities.

- **Output validation and lack of standardized benchmarks:** Unlike text and image models, where output can be immediately scored, measuring the quality of biological foundation models requires wet-lab validation. This makes iterating much slower and cost-intensive than in other verticals. Additionally, the lack of standardized benchmarks across models presents a challenge in quantifying the scale of progression and mapping the rapidly growing field of new models.

- **Workflow integration:** Biopharma R&D is inherently complex, built upon entrenched processes and optimized for human-led discovery and decision-making. Effective integration of AI will require not only technical capability but also organizational management.

- **Biosecurity:** While still early, the ability of new genomic foundation models to optimize or generate novel viral genomes presents a significant biodefense risk. As these models' capabilities improve, equally powerful tools will be required to detect and prevent viral threats. Underscoring this need, AI-enabled biosecurity platform Valthos raised a $30 million Series A from OpenAI, Lux Capital, Founders Fund, and Definition in October.

Outlook

Domain-specific models have already established themselves as valuable computation tools, providing tangible savings by shifting early screening, validation, and target identification in silico. The value of larger generalized foundation models and emerging AI-native biotech startups is far less established but is entering a crucial period of validation. With several AI-discovered drugs currently in clinical trials and leading startups such as Xaira Therapeutics and Isomorphic Labs likely to enter the clinic within the coming years, the clinical efficacy of AI-discovered assets will become clearer. While AI tools for experimental design and clinical-trial optimization will likely continue to be integrated into drug development, the success of AI-discovered assets will be the best benchmark for the technology and will have the most significant impact on the biopharma business model.