HSBC Global Investment Research

Future Healthcare
Equity Research Report
China | Thematic research

By: Linda Shu (S1700522120001), Andre Sun (S1700124030009)

July 2026 www.research.hsbc.com

SPOTLIGHT

China Healthcare: AI in Drug Discovery (AIDD)

From experimentation to monetisation

We provide a roadmap for navigating the AI drug discovery sector, one of China's most narrative-driven and volatile sectors

Valuations are powered more by optionality than fundamentals, the biggest opportunities stem from market mispricing and volatility

In accompanying reports, also published today, we initiate on Insilico Medicine at Buy and XtalPi at Hold



No country for bears

The 24th edition of the EM Sentiment Survey

Click to view

Disclosures & Disclaimer: This report must be read with the disclosures and the analyst certifications in the Disclosure appendix, and with the Disclaimer, which forms part of it.

Why read this report?

- We provide a roadmap for navigating the AI drug discovery sector, one of China's most narrative-driven and volatile sectors
- Valuations are powered more by optionality than fundamentals, the biggest opportunities stem from market mispricing and volatility
- In accompanying reports, also published today, we initiate on Insilico Medicine at Buy and XtalPi at Hold

A transformative sector, but still no consensus

Using AI to discover new drugs is rapidly becoming one of China's largest and most misunderstood investment themes. Billions of dollars are flowing into the sector, partnerships with global pharma names are accelerating, and valuations are increasingly pricing in the potential of a company's platform, rather than current earnings. Yet despite the excitement, investors still face one critical problem: there is no consensus on what these AI drug discovery (AIDD) companies are – software, biotech, contract research organisations (CROs) that conduct research and development for drug companies, or future platform monopolies?

This uncertainty is driving one of the biggest valuation dispersions in healthcare and AI

We highlight that on one side, AI is clearly transforming drug discovery workflows, compressing timelines, reducing experimental costs, and reshaping how pharma allocates R&D. On the other, monetisation remains early, business models are still evolving, and many public names are trading more on narrative, optionality, and AI scarcity premium than predictable cash flows.

Given the potential upside but also significant mispricing risk, we address the five biggest investor debates shaping the AIDD sector today:

1. **Where is monetisation emerging? Which business models are structurally durable vs vulnerable to commoditisation?**
2. **What actually moves these stocks?**
3. **How should investors value platform optionality?**
4. **What are long-term moats in a world where AI models can rapidly be commoditised?**
5. **Which parts of the value chain capture economics – and which get compressed?**

Our differentiated approach combines five elements:

We believe the market is using outdated frameworks to analyse this rapidly changing industry

- US vs China strategic comparisons and competitive positioning.
- Lessons from the early internet platform evolution and CRO industry consolidation.
- A framework for identifying where future pricing power and data moats emerge.
- Analysis of where valuation methodologies are likely wrong – particularly around platform scalability, talent value, and downstream asset ownership.
- A roadmap for navigating one of the market's most narrative-driven and volatile sectors.

This is our latest report on the Future Healthcare theme. If you want to subscribe to any of our nine big themes, *click here*.

Our key takeaway? The winners in AIDD may not be the companies with the best AI models today, but the platforms that can keep learning from their own data, fit deeply into pharma workflows, and repeatedly create valuable drug assets at scale.

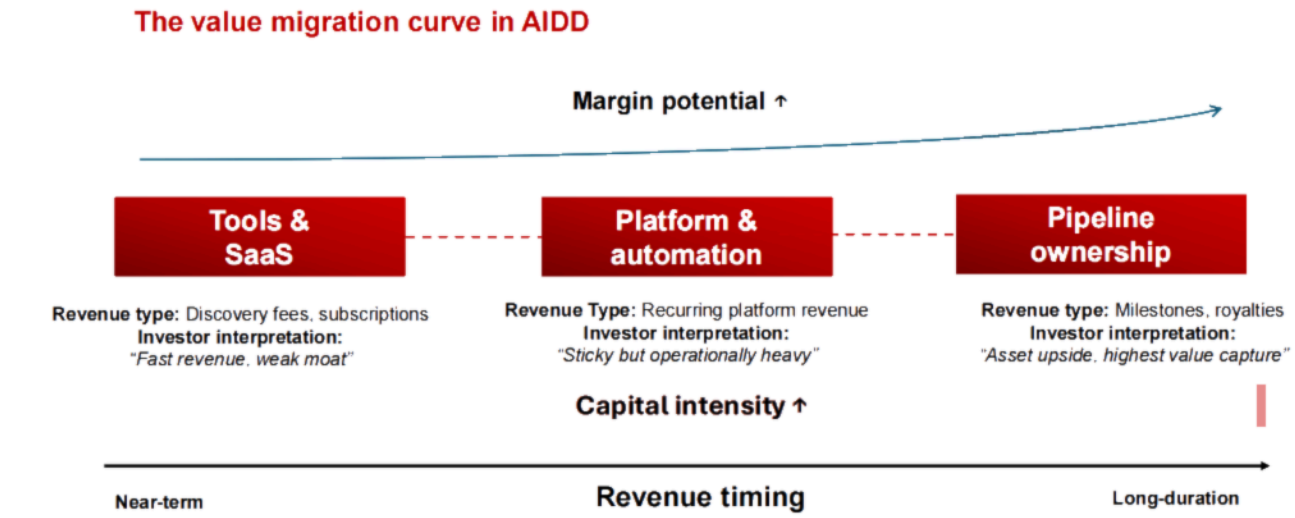
This sector could become one of the defining healthcare platform transitions over the next three-to-five years. But separating durable platform winners from temporary AI hype will require a fundamentally different investment framework, which is the the subject of this report.

Key investor debates: Our core findings

1. **Where is AIDD monetising today? Which business models are durable vs at risk?** Current monetisation is real but still concentrated in lower-value layers like software as a service (SaaS) offerings, drug discovery fees, and CRO services. We see long-term value shifting downstream towards asset ownership, proprietary data, and improving platform economics. Pure SaaS providers face rising commoditisation risk.
2. **What moves AIDD stocks − and why does leadership change so quickly?** AIDD trades on three drivers: narratives, event catalysts, and fundamentals. In bull markets, the AI narrative typically dominates. In risk-off periods, cash runways and monetisation matter a lot more. Timing when these factors rotate and become more important is critical.
3. **What is the correct valuation framework?** We believe traditional DCF frameworks are insufficient as AIDD companies increasingly trade as "stacks of options" − combining platform capability, talent scarcity, proprietary data, and future ecosystem value. Markets are pricing their future potential trajectory, not their current cash flow.
4. **What becomes the long-term moat?** We expect most AI advantages to become commoditised. Instead, durable moats will emerge from propietary data generation, workflow integration, regulatory positioning, and wet-lab infrastructure − not simply better algorithms.
5. **What lessons can investors learn from the internet and CRO industries?** Like the early internet and CRO leaders, AIDD winners may evolve from fragmented tools into integrated platforms. Historical winners captured ecosystems, embedded workflows, and long-term switching costs − not feature-level innovation alone.
6. **What becomes commoditised vs defensible in AI drug discovery?** Generalised modeling, protein folding, and molecular design capabilities are rapidly being standardised. Defensible value increasingly sits in proprietary biological data, patient access, experimental feedback loops, and scalable asset-generation infrastructure.

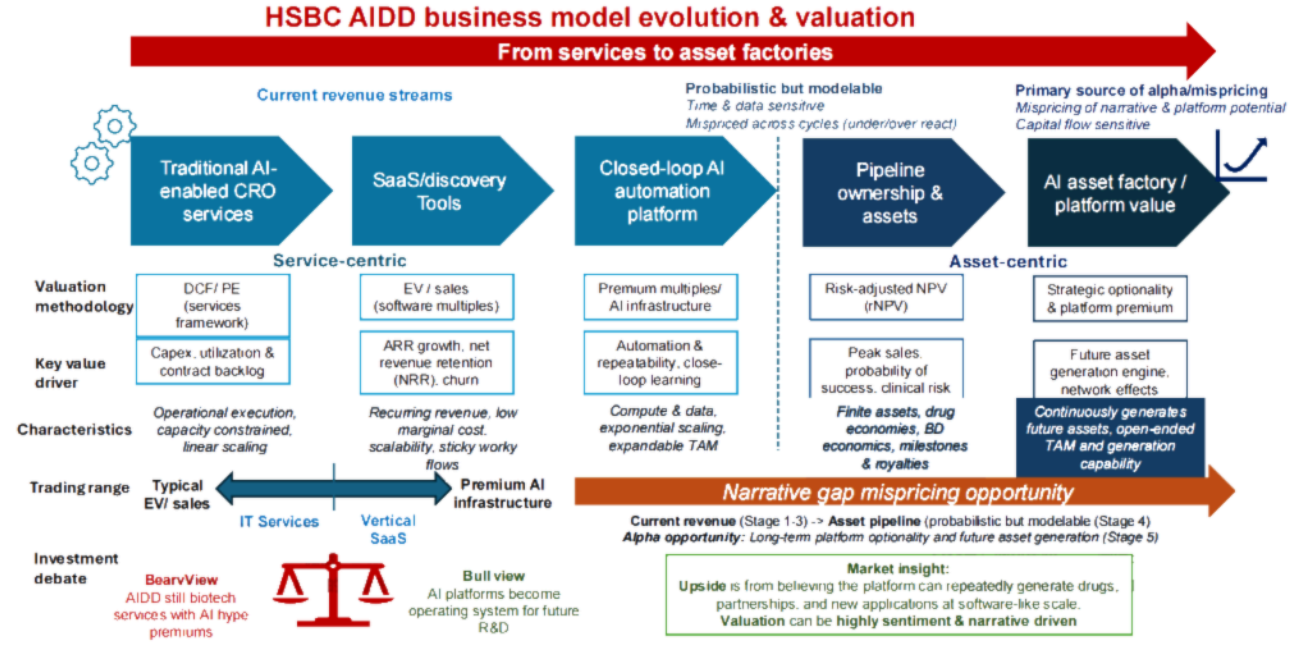
Our investment thesis in exhibits

Exhibit 1. AIDD: The value migration curve from tools to pipeline ownership



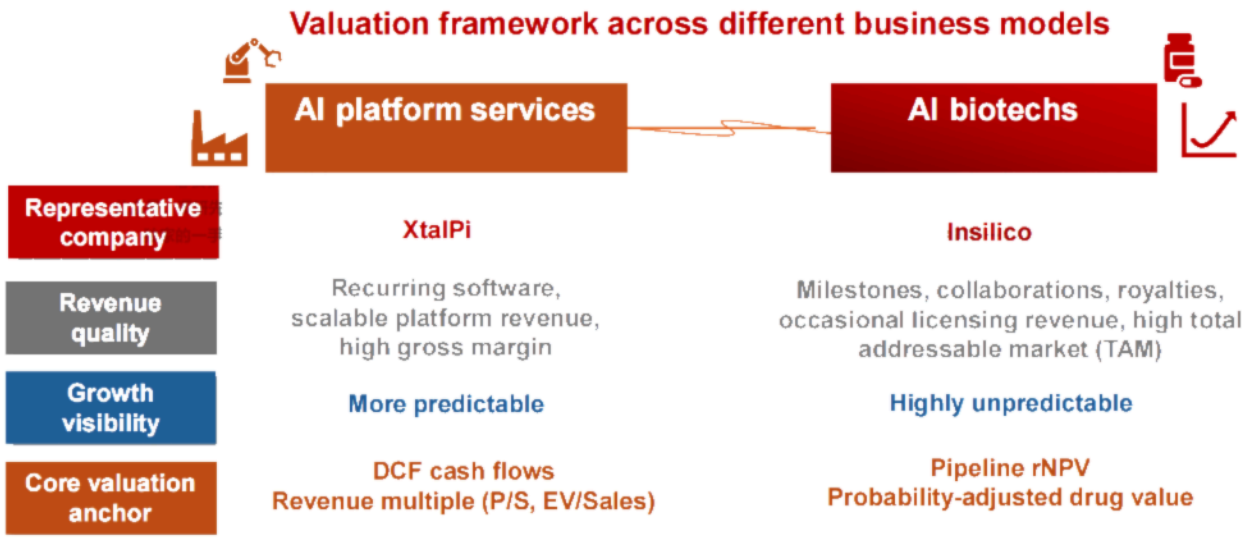
Source: Pharmcube, Company reports, HSBC Qianhai Securities

Exhibit 2. HSBC AI healthcare valuation framework: "Trading the Stack"



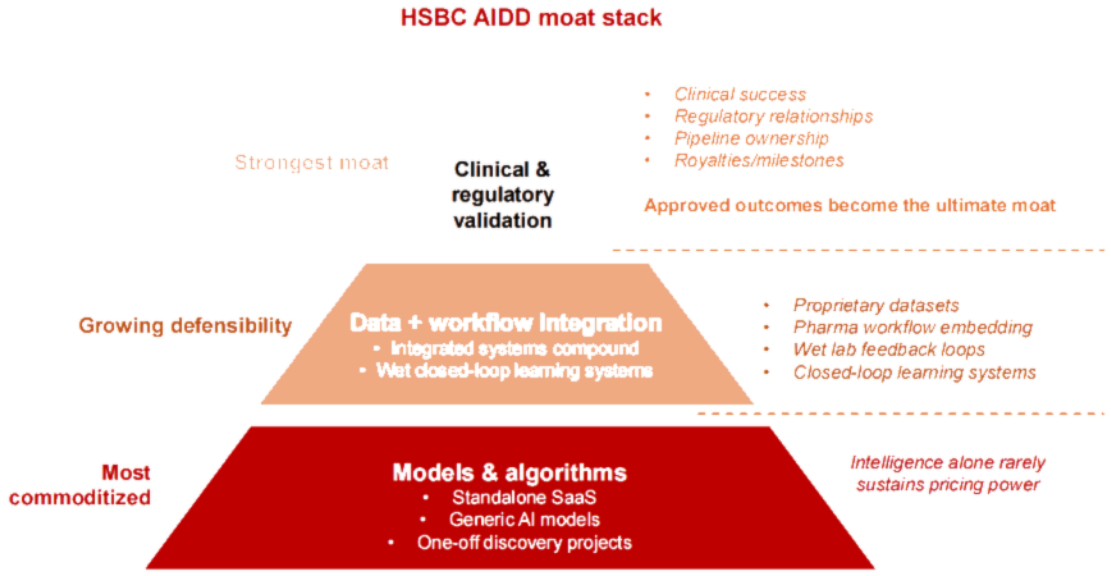
Source: Pharmcube, Company reports, HSBC Qianhai Securities

Exhibit 3. Our valuation framework across different AI and healthcare business models



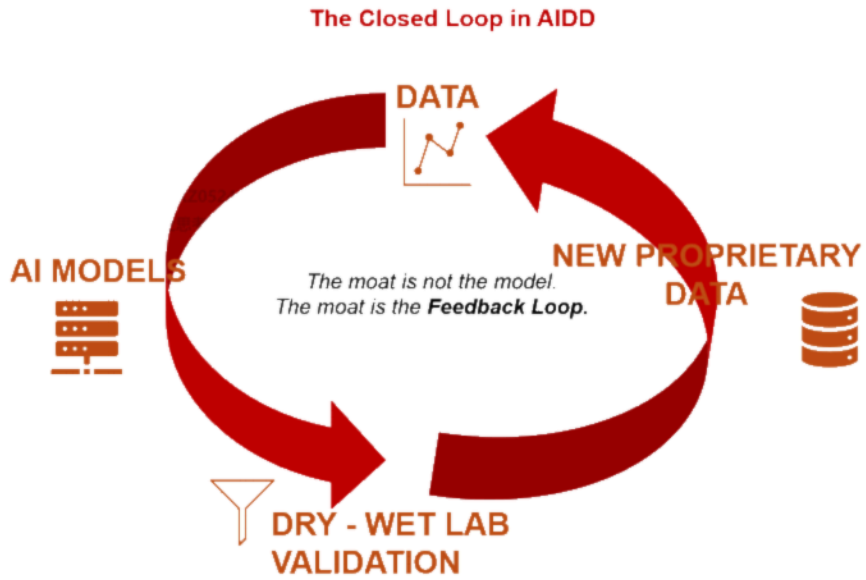
Source: Pharmcube, Company reports, HSBC Qianhai Securities

Exhibit 4. AI healthcare moat framework: Separating signal from noise



Source: Pharmcube, Company reports, HSBC Qianhai Securities

Exhibit 5. The moat is not the model it's the feedback loop



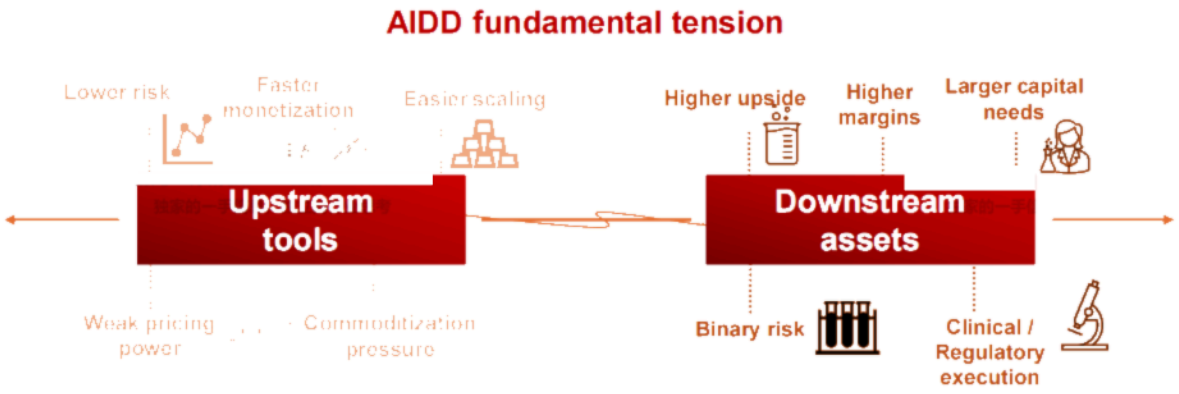
Source: Pharmcube, Company reports, HSBC Qianhai Securities

Exhibit 6. Global AIDD market trend: US and China are building ecosystems with different approaches



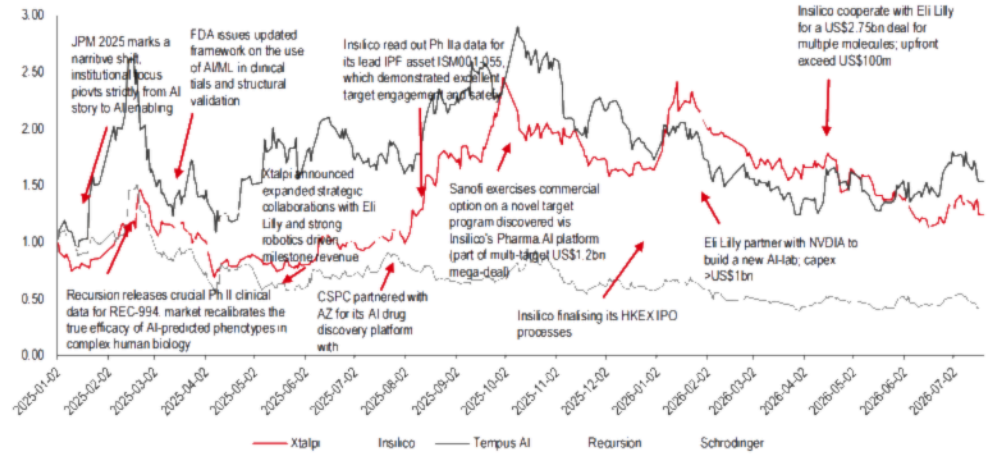
Source: Pharmcube, Company reports, HSBC Qianhai Securities

Exhibit 7. Tensions in AIDD – moving downstream (higher margins, higher risk, higher capital needs) vs upstream (lower risk, lower returns, commoditisation pressure) determines how business models evolve



Source: Pharmcube, Company reports, HSBC Qianhai Securities

Exhibit 8. Sector share price rally with major BD announcements, clinical data readouts and regulatory tailwinds



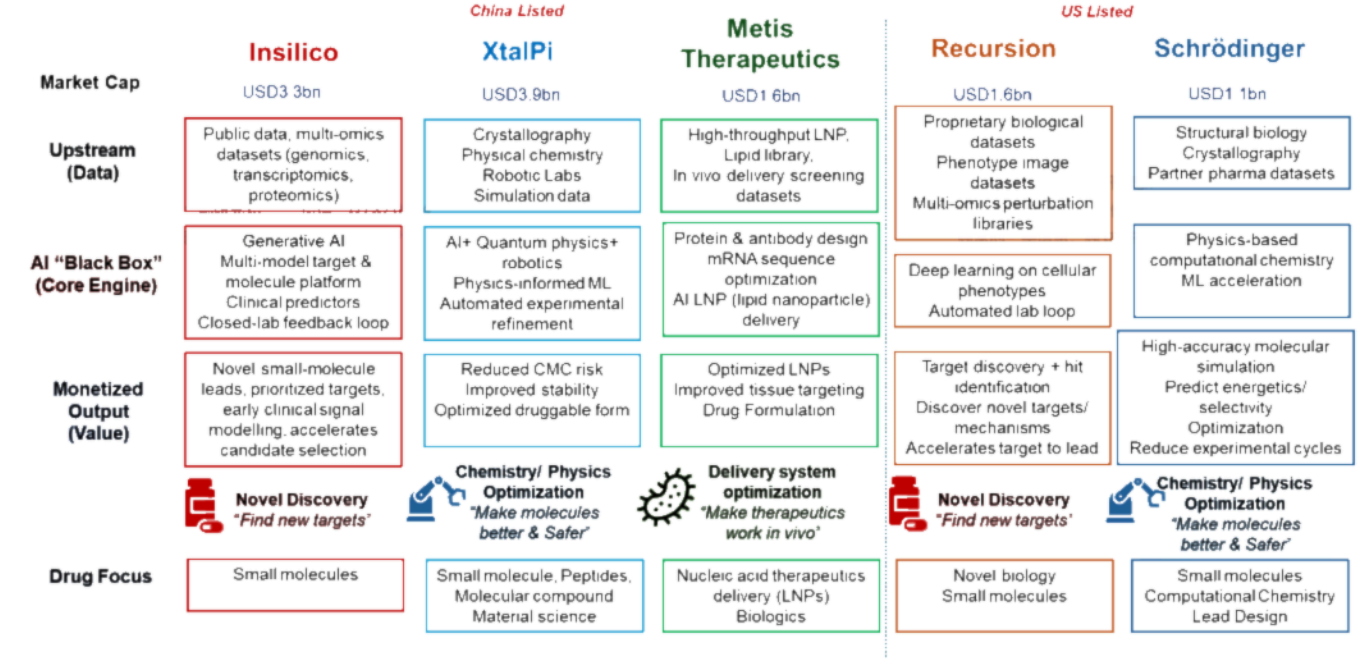
Source: Company reports, HSBC Qianhai Securities

Exhibit 9. Three drivers of the AI healthcare sector: Fundamentals, events, and narrative; the market often reprices on sentiment or correlation shifts

Phase	Driver in control	What moves stocks?
AI rally	Narrative	Tech multiples, flows
Mid-cycle	Events	Deals, clinical updates, AI progress
First crack	Narrative (down)	Macro, liquidity
Drawdown	Fundamentals	Cash, backlog

Source: HSBC Qianhai Securities

Exhibit 10. Technological comparison of major players in AI + healthcare



Source: Company data, HSBC Qianhai Securities

Related research

Recommended reading...

- *China Healthcare – AI in drug discovery: From experimentation to monetisation*, Linda Shu, **22 July 2026**
- *Insilico Medicine – Initiation at Buy: From concept stock to solid execution*, Linda Shu, **22 July 2026**
- *Xtalpi – Initiation at Hold: China's first profitable AIDD platform*, Linda Shu, **22 July 2026**
- *Spotlight: China Healthcare – A roadmap to how AI can transform the sector*, Linda Shu, **5 March 2025**
- *AI and China Healthcare – Turning innovation into income: how AI is powering growth*, Linda Shu, **27 February 2026**
- *China Healthcare: 2026 outlook: From business development deals to global delivery, soft recovery in domestic market*, Linda Shu, **6 January 2026**
- *CSPC – Buy: USD18.5bn peptide platform deal reached*, Linda Shu, **4 February 2026**
- *United Imaging – Buy: Breakthrough 'videography' MRI product launched*, Linda Shu, **4 February 2026**
- *China Biotechnology: Going global: China remains the global innovation hub in 2025*, Cindy Chai, **14 February 2025**
- *China Healthcare: Future healthcare: Embracing commercial medical insurance*, Linda Shu, **22 January 2025**
- *Spotlight: Quantum Computing Age: Multi-sector disruption*, Davey Jose (HSBC Global Investment Research), **5 October 2022**
- *Spotlight: Surgical Robots: At the cutting edge of growth*, Helen Fang (HSBC Global Investment Research), **17 November 2023**
- *Generative AI: The rise of the machines*, Mark McDonald (HSBC Global Investment Research), **23 February 2023**

Five key investor debates

1. Where is AIDD monetising today? Which business models are durable vs at risk?

A. Cash flow visible, but not yet at scale or predictable

AIDD is generating revenue today, but still not at levels where investors would like it. Despite major progress having been made, most of current monetisation sits in the low- to mid-value layers: SaaS tools, discovery collaborations, and fee-for-service modeling. These produce visible but unstable revenue, and are often tied to early-stage research.

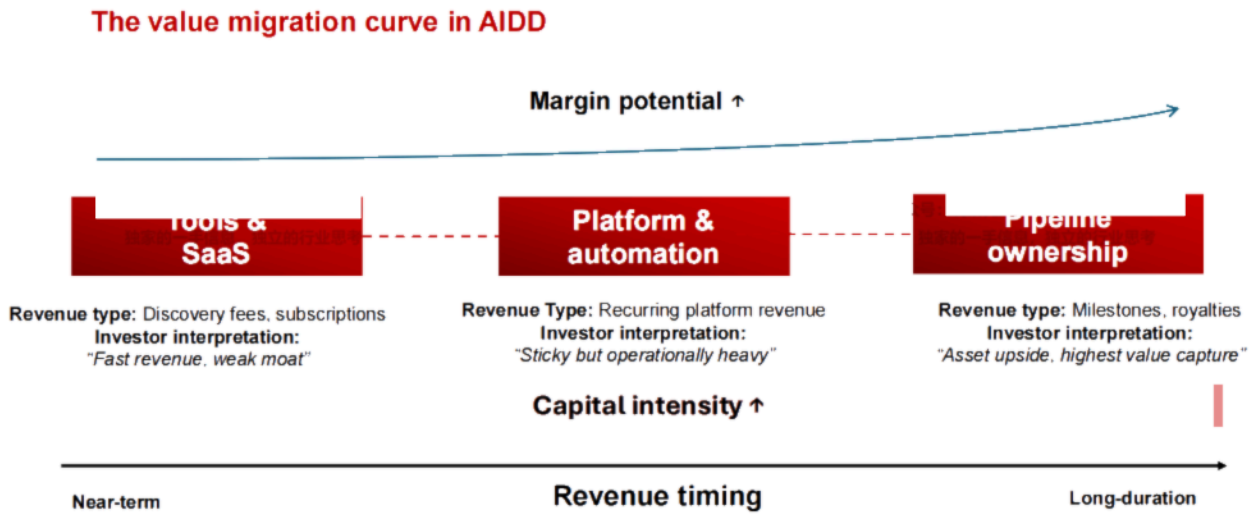
Investor pain point remains highly speculative as the industry is still in its early development phase: is this software (high-margin, scalable) or biotech (high-risk, capital-heavy), or CRO (platform services)? There's no clear consensus from the street in terms of which valuation to use or what business model to classify AIDDs as given these companies are also still navigating themselves the fast-changing world of AI. Until companies prove repeatable SaaS scale or clinical asset realisation, we believe multiples could remain unstable.

Future pivot: From selling picks to owning gold – services & SaaS are entry points, not endpoints. The revenue durability question hinges on what gets commoditised vs what compounds. General AI capabilities like modeling, protein folding, and molecular design are rapidly being standardised. As Google DeepMind-type foundational players push performance boundaries, pure-play SaaS AIDD is becoming price-taker infrastructure, not a value owner.

However, we identify where durability is emerging to differentiate the potential "winners" in the vertical applications of AI in healthcare:

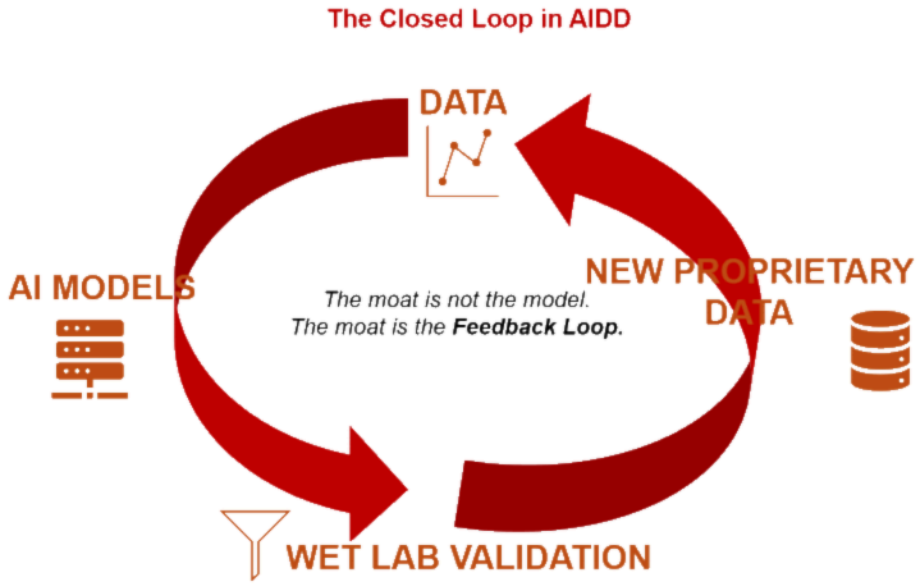
- **Closed-loop platforms and infrastructure** (AI + wet lab + proprietary data): Harder to replicate, improving with scale.
- **Exclusive data moats/patient-level data access:** Structural moat in a regulated environment (e.g. US FDA constraints).
- **Downstream economics participation/pipeline ownership:** Milestones, royalties, equity in assets.

Exhibit 11. AIDD: The value migration curve from tools to pipeline ownership



Source: Pharmcube, Company reports, HSBC Qianhai Securities

Exhibit 12. The moat is not the model it's the feedback loop



Source: Pharmcube, Company reports, HSBC Qianhai Securities

B. Market trend – booming deal activities: Commanding valuation reflects future platform potential instead of current revenue

Exhibit 13. Global AIDD market trend: US and China are building ecosystems with a different approach



Source: Pharmcube, Company reports, HSBC Qianhai Securities

2026 y-t-d deal flow shows a clear pattern: Capital is chasing capability, not cash flow. Most AIDD platforms remain pre-profit, yet command premium valuations. Valuations are based on future platform output, not current revenue.

Overall, we find the following to be highly valued in primary markets:

People: Founders, ex-pharma leaders, top AI researchers.

Data advantage: Exclusive partnerships, proprietary clinical, and longitudinal datasets.

Platform throughput: How many drug candidates can be generated per year.

Early validation: Investigational new drug (IND) filings, pharma partnerships.

Though there is no standard metric, investors triangulate on implied value per pipeline asset, platform replacement cost (data + compute + talent), and strategic premium from pharma optionality.

The key market takeaway? *The premium valuations are a function of future pipeline optionality, as in the platform's ability to generate future assets, not present earnings, which means investors risk paying high multiples without proof the platforms will perform well. However, investors are speeding up their entry into this space to gain the "optionality" of potentially significant upside, given biology AI is much harder than scaling general large language models (LLMs) and has much higher barriers to entry.*

Exhibit 14. Major deals in AIDD (China) 2026 y-t-d

Company/Acquirer	Counterparty	Platform focus	Deal value	Why it matters	Implied valuation logic
Pharma / biotech platform consolidation in AIDD; AI-biopharma ecosystem is supporting biotechs in "Going Global"					
Eli Lilly	Insilico Medicine	End-to-end generative-AI drug platform	Up to USD2.75bn (USD115m upfront + milestones + royalties; global rights to preclinical assets)	One of the major AIDD platform deals to date	The market now rewards AI platforms with clinical validation rather than pure software claims. Insilico's valuation is increasingly benchmarked to biotech pipelines, not SaaS companies.
Insilico Medicine	Servier	AI oncology discovery platform	Up to USD888m	One of the largest Europe-China AI-drug licensing deals	This validates that pharma now treats Chinese AI-originated molecules as globally licensable assets rather than "China-only" innovation. The upfront payment was only cUSD32m, again reinforcing milestone-heavy biotech-option pricing.
Insilico Medicine	Qilu Pharma	Cardiometabolic small molecules via Pharma.AI	Near USD120m	Significant because Qilu moved from software customer to co-development partner	This suggests China pharmas increasingly want vertically-integrated AI discovery capability rather than external software tools. Comparable to how cloud customers evolved into strategic AI co-builders.
XtalPi	Pharma collaborations	AI chemistry + robotic labs	Mostly undisclosed	Strong China biopharma penetration	The China market values integrated robotics + chemistry execution capabilities more than standalone frontier models.
XtalPi	Sunshine Lake Pharma	AI-driven automated drug R&D JV	Undisclosed	Joint AI lab + Design-Build-Test-Learn closed loop	XtalPi is evolving from "AI chemistry software" into an industrialised autonomous R&D infrastructure provider. The market likely values robotics + the wet-lab automation premium more than pure AI multiples.
Large China pharmas (Hengrui / CSPC / Fosun ecosystem)	Multiple AI startups	AI-enabled molecule generation + licensing	Mostly undisclosed	Sharp increase in outbound licensing activity	China pharmas are now using AI platforms to generate exportable preclinical assets aimed at US/EU out-licensing economics.
China tech expanding into regulated healthcare workloads and healthcare infrastructure					
Alibaba Group cloud ecosystem	China healthcare AI ecosystem partnerships	Biomedical foundation models	Infrastructure-scale investment	Focused more on compute + ecosystem enablement than acquisitions	China techs appear to be pursuing "AI utility layer" economics: monetise inference + cloud + healthcare deployment rather than owning biotech risk directly.
Tencent ecosystem	Hospital / cloud / healthcare AI partnerships	Clinical AI + healthcare infrastructure	Mostly undisclosed	Increased AI capex and healthcare cloud push	Tencent's strategy resembles an enterprise healthcare operating system expansion rather than molecule ownership.

Source: Company data, HSBC Qianhai Securities

Major market trends:

- Frontier AI and Big Tech are moving into healthcare and biology verticals **(US and China).**
- Rising pharma and biotech platform collaborations in AIDD **(US and China).**
- **Chinese AI-biopharma ecosystem platforms are increasingly exporting R&D/IP, and are not just service vendors anymore**: Large pharma increasingly accepts China-originated AI molecules into global pipelines, with commercial models converging toward US biotech economics (upfront + milestones). This is supported by the Eli Lilly-Insilico Medicine deal (cUSD2.75bn). China's unique **"AI-pharma/CRO" hybrid model** could become a competitive advantage.
- **China's big tech strategy differs from the US:** More ecosystem/platform enablement, hospital digitisation/workflows, and AI infrastructure rather than outright biotech acquisitions.
- **US emphasises** frontier foundation models, proprietary multimodal biological data, and agents; **China emphasises** scalable design-build-test-learn execution, throughput economics, AI & robotics, and automated wet-lab integration.

Exhibit 15. Major deals in AIDD (US) 2026 y-t-d

Company/Acquirer	Target/Partner	Platform focus	Deal value	Notes	Implied valuation/strategic logic
Big pharma/biotech platform consolidation: Pharma buys pipelines and data, not just models					
Schrödinger + Eli Lilly	TuneLab integration into LiveDesign	AI-native molecular design stack	Undisclosed	Eli Lilly is taking its specialised, proprietary AI software it built and making it available to external biotech startups (externalising).	Shifts the AI moat from "models" to "workflow distribution." Pharmas increasingly values interoperable AI operating systems rather than standalone discovery tools.
Tempus AI	Ongoing oncology data / AI partnerships	Precision medicine + multimodal data	Multiple undisclosed deals	Tempus increasingly positioned as a healthcare data layer.	Valuation logic resembles a "healthcare Bloomberg": premium assigned for longitudinal clinical data network effects.
Recursion Pharmaceuticals	Multiple AI-enabled pipeline partnerships	Phenomics + automated biology	Various	Recursion continuing its consolidation strategy around industrialised wet-lab automation.	Investors increasingly underwriting Recursion as an "AI factory" rather than a single-asset biotech.
Schrödinger	Enterprise pharma integrations	Physics + ML molecular simulation	Various	Strong positioning as a picks-and-shovels layer.	Valuation supported by hybrid software + royalties model, which compresses downside vs binary biotech risk.
Big techs are expanding from horizontal models to verticalised healthcare					
Anthropic	Coefficient Bio (reported acquisition)	Foundation models for biology / autonomous science	Reported cUSD400m Source: Fierce Biotech	Team folded into healthcare/life sciences initiative.	Early evidence frontier model labs are buying domain-native scientific-agent teams rather than building internally. Valuation appears to include a talent + proprietary wet-lab data premium
Anthropic	Claude for Healthcare rollout	Clinical workflow copilots	N/A (product expansion)	Healthcare-specific enterprise stack.	Strategic move mirrors cloud wars: captures regulated enterprise workflows first, monetise later via API + agent orchestration.
OpenAI	Multiple healthcare/research integrations (non-M&A)	Scientific reasoning / biomedical copilots	Mostly undisclosed	Focus increasingly shifting from generic LLMs to domain-specific agents.	OpenAI's healthcare push is being valued indirectly through enterprise ARR multiples rather than standalone healthcare economics; biology seen as a major future inference-demand vertical.
AI infrastructure: Supporting compute power and integration with clinical, workflow platforms					
Eli Lilly + NVIDIA	AI discovery infrastructure collaboration	BioNeMo + accelerated drug design	Up to USD1bn co-investment	AI compute becoming strategic pharma infrastructure.	Suggests future differentiation may depend on proprietary compute + multimodal biological data, not only algorithms.

Source: Company data, HSBC Qianhai Securities

C. Business model evolution: What's the tension?

AIDD business models are converging but not all paths will lead to value, in our view. We identify three monetisation layers that exist today:

(1) **Entry layer: Tools/SaaS/Discovery fees:** This is the fastest to revenue, has the lowest barriers and usually used in preclinical stages. The core problem is limited scale, a high commoditisation risk, and weaker pricing power.

(2) **Mid layer: CRO/Platform services:** This combines AI + lab execution and generates real, recurring revenue. The tradeoff is its operationally heavy and caps margins.

(3) **Value layer: Pipeline ownership/Licensing: Revenue via BD, milestones, royalties:** We see this with the highest upside but is highly volatile and back-loaded.

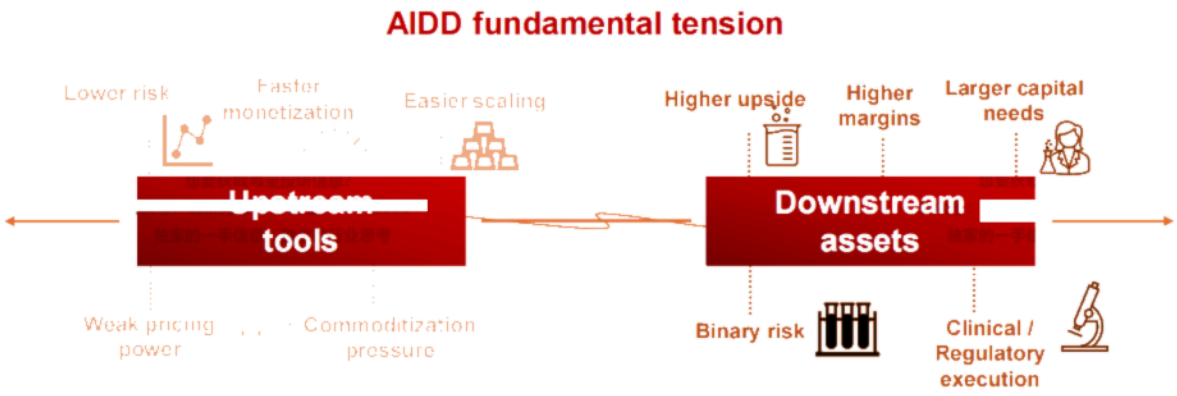
We see that companies are shifting decisively towards asset-centric models, as tools don't capture the drug value like assets do. But this also introduces a fundamental tension:

Moving downstream = higher margins, higher risk, higher capital needs.

Staying upstream = lower risk, lower returns, commoditisation pressure.

Ultimately, we see the winners ***emerging as hybrid operators***, using services revenue to fund internal pipelines, while building data engines and scale into multi-asset portfolios with diversified revenue streams. **The winners** are likely to evolve as platforms that can consistently originate and monetise assets (instead of pure tools and services).

Exhibit 16. AIDD business model evolution: Fundamental tension



Source: Pharmcube, Company reports, HSBC Qianhai Securities

2. What moves stocks (and why do the drivers change rapidly)?

The mistake most investors make is treating AI healthcare as a single-factor growth story. We don't believe it is. Instead, we view it as a three-factor trade with unstable leadership, **where returns rotate between fundamentals, events, and the narrative** – and they can rotate faster than most portfolios can adjust. At any given time, one of these factors dominates more. The edge comes with timing the rotation, or inflection points.

The three drivers in the AI healthcare trade:

1. **Fundamentals set the floor:**

In theory, recurring platform revenue and pharma partnerships should anchor valuations. In practice, most "revenue" in AI healthcare is milestone-based, back-ended, and only loosely predictive of future cash flow. Partnership backlogs provide validation, but they are not the same as visibility.

This is why fundamentals can sometimes be largely ignored in rising markets. They only become relevant when liquidity tightens and investors are forced to ask a much simpler question: *who survives?* At that point, the cash runway and deal coverage abruptly replace TAM narratives.

2. **Events create the spikes and stock differentiation:**

AI does not remove the uncertainty from clinical trials, but accelerates exposure to it. Clinical milestones, IND filings, and pharma deals still drive step-function moves, often with classic biotech asymmetry risk-reward: sharp rallies on positive data and deep drawdowns on failure. What AI changes is it could amplify the frequency and perception of these events. As a result, events generate alpha, but they are not a stable driver of sector-wide performance.

This also matters most in the middle of the cycle, when the narrative weakens and investors start discriminating.

3. **Narrative beta and platform capability drive most upside and are least modeled:**

In bull markets, AI healthcare stocks trade less like biotech names and more like long-duration tech stocks. Multiples expand with the broader AI trade, capital rotates into "next wave" themes, and a scarcity premium emerges around pure-play names.

This creates a powerful dynamic – stocks can move without company-specific news. Correlations shift away from clinical timelines and toward AI sentiment, tech multiples, and liquidity conditions.

This is why the biggest rallies in the space often occur *before* revenue or clinical validation. And it's also why those gains can unwind just as quickly when the AI trade softens.

What dominates and when?

In risk-on environments, narratives can drive most of the returns. Investors are effectively paying for exposure to AI rather than healthcare outcomes. Valuations expand, dispersion compresses, and even early-stage names re-rate aggressively. Fundamentals are acknowledged but not priced.

As markets transition, the narrative loses momentum, dispersion rises, and investors begin to differentiate between companies with real partnerships or pipelines and those relying more on a story. This is where event-driven performance becomes more important.

In risk-off regimes, the hierarchy of drivers flips. Fundamentals take control, the narrative loses momentum, and capital retreats to names with balance sheet strength and visible deal flow.

Implications for investors: Timing the rotation

What makes this sector difficult and tradable is not just the existence of these three drivers, but the speed at which the shift from narrative to fundamentals happens, driven by flows and macro conditions.

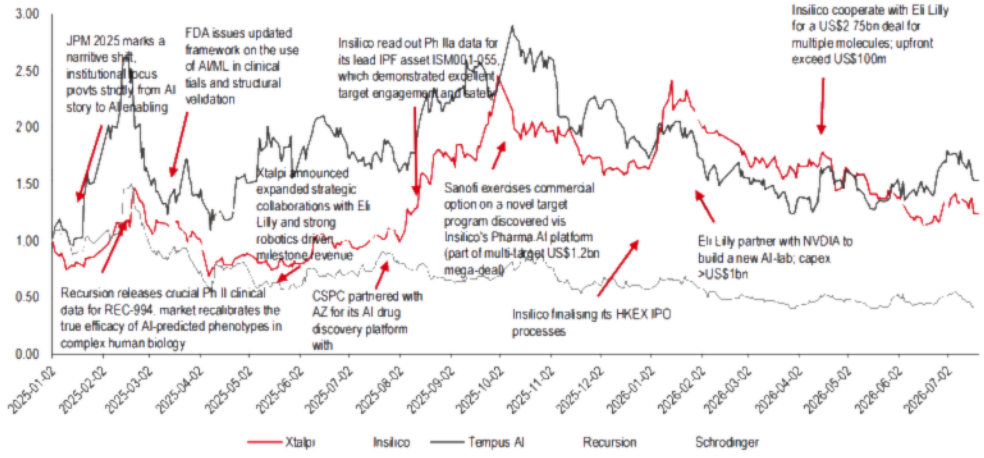
In bull markets, investors are effectively long AI beta with biotech optionality. In drawdowns, investors risk holding pre-revenue biotech with compressed tech multiples. The underlying companies have not changed, but the driver of their valuation has.

Exhibit 17. Three drivers of AI healthcare trade: Fundamentals, events, and narrative; the market often reprices on sentiment or correlation shifts

Phase	Drivers in control	What moves stocks?
AI rally	Narrative	Tech multiples, flows
Mid-cycle	Events	Deals, clinical updates, AI progress
First crack	Narrative (down)	Macro, liquidity
Drawdown	Fundamentals	Cash, backlog

Source: HSBC Qianhai Securities

Exhibit 18. Sector share price rally with major BD announcements, clinical data readouts, and regulatory tailwinds



Source: Company reports, HSBC Qianhai Securities

3. What is the correct valuation framework when the investor base is gaining optionality, not cash flows?

Healthcare is becoming a primary vertical battleground with sizeable TAM, high-value data, and unsolved problems in drug discovery. However, investors are speeding up their entry into this space to gain the "optionality" of potentially significant upside, given biology AI is much harder than scaling general large language models (LLMs) and has much higher barriers to entry.

Trading optionality, not cash flows:
AI healthcare companies are not single assets, but more like a stack of options. As we believe the traditional DCF approach is structurally insufficient, we breakdown valuations into three layers − what part of the valuation is real (i.e. based on tangible results), what part is probable (i.e. expected), and what part is story driven − and which one is currently mispriced?

Case study: Anthropic acquisition of Coefficient Bio (cUSD400m, c0.1% of Anthropic equity). This transaction (<10 employees, <1 year old, stealth/no revenue or pipeline) reveals an important implication that valuation methodologies do not always follow a traditional framework. The market assigns material value to "**future platform capability and trajectory**", well before monetisation exists.

What is actually being valued?

- **Talent scarcity:** Top-tier AI + biology talent = scarce asset (cUSD40-50m per researcher).
- **Time-to-market advantage:** Pre-built trajectory with a team already aligned on biological modeling architectures and experimental design workflows.
- **Strategic entry into biology:** Proprietary biological data and domain-specific expertise becomes a key moat, while general LLMs are commoditised.
- **Platform integration:** Anthropic can integrate Claude models with biology-specific vertical integration, creating a full-stack scientific AI platform.

What justifies a significant valuation and what is the valuation logic?

- **Replacement cost/time saved framework:** How much does it cost and time consumed to build internally?
- **Venture-style "power law" optionality:** If successful, a platform could unlock multi-billion dollar upside via a drug discovery platform and pharma partnerships (significant TAM vertical).
- **Data and moat premium:** Biological datasets are scarce, proprietary and cumulative – owning a team that can generate novel biological data loops is extremely valuable.
- **Data assets create future monetisation value:** e.g.: electronic health record (EHR) platforms and genomics datasets.
- **Platform multiple or potential:** Long-term platform option.
- **Talent and strategic option value:** Access to an elite team and entry into biotech AI.

Where the market could be wrong?

Mispricing 1: Platform optionality and ecosystem expansion potential

Investors could underestimate repeatable asset-factory monetisation and partnership generation capability, particularly when it evolves faster than originally perceived.

Mispricing 2: AI scalability could be under-modeled

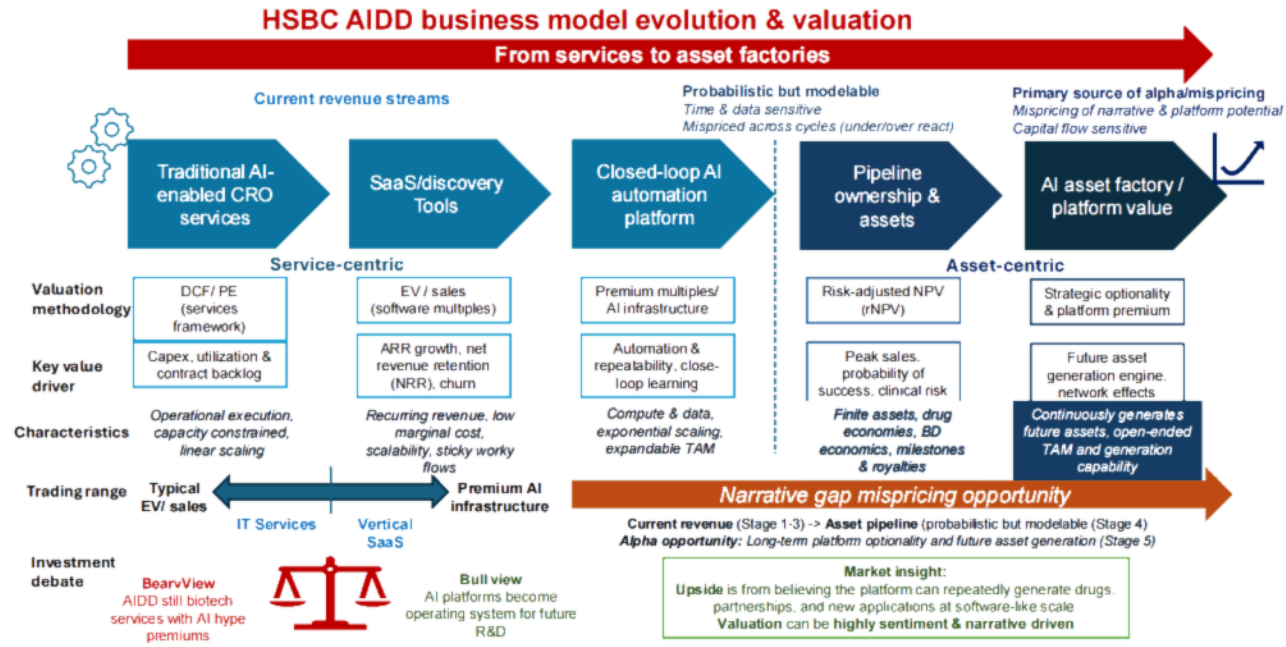
The market could model AI drug discovery as a linear biotech, while in reality it ignores scalable platform infrastructure economics.

Mispricing 3: Talent is not capitalised properly

The market still expenses R&D, while frontier deals imply talent should also be treated as a balance sheet asset.

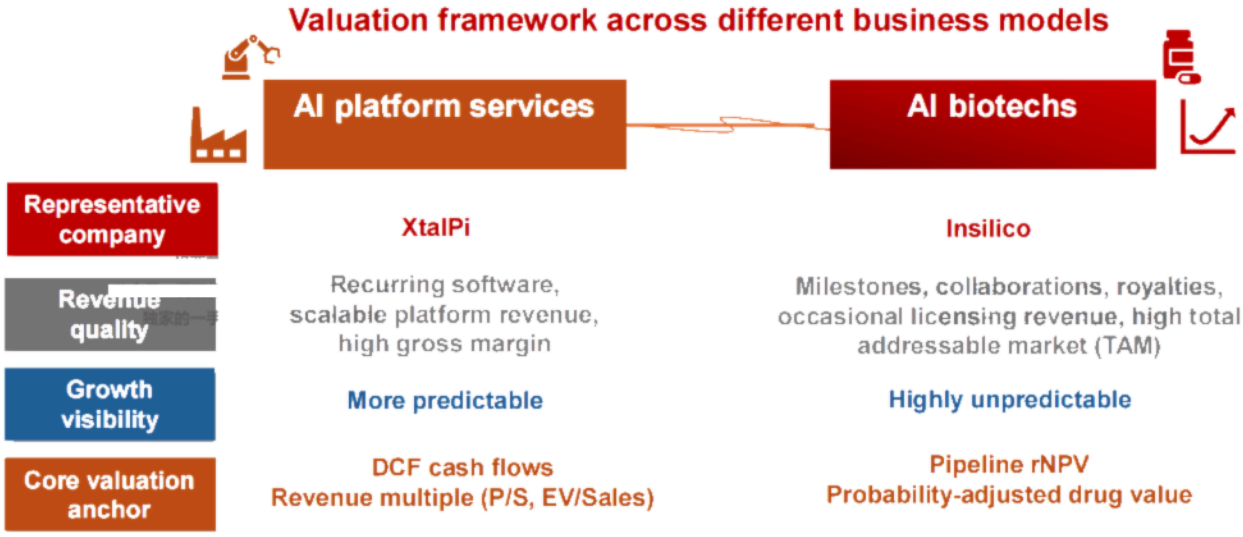
HSBC AI Healthcare Valuation Framework: "Trading the Stack"

Exhibit 19. HSBC AI Healthcare valuation framework: "Trading the Stack"



Source: Pharmcube, Company reports, HSBC Qianhai Securities

Exhibit 20. Our valuation framework across different AI + healthcare business models



Source: Pharmcube, Company reports, HSBC Qianhai Securities

4. What is the long-term business moat? What separates "platform winners" from "commoditised AI"?

This is one of the most debated and speculative questions for investors today. We reframe this, so instead of debating "Who has the best AI?", we believe the more appropriate question to ask is: *Where does long-term pricing power actually come from, and which AI advantages will inevitably be commoditised?*

AI Healthcare: Separating durable moats from temporary "AI advantages"
In our observation and view, most of today's "AI edge" is transient. The industry is seeing an oversupply of models, tools, and talent, driving low switching costs and rapid capability diffusion.

The real differentiation will not be *who builds AI*, but **who embeds AI into proprietary workflows, data loops, and clinical outcomes that are hard to replicate**.

We also study two case studies from previous development cycles from the early Internet and CRO, helping us to assess where true moats could emerge in AI and healthcare.

AI healthcare's industry current development stage echoes two historical playbooks: early Internet + CRO

AI healthcare is at a familiar **early-stage inflection point:**

- High growth, high capex, rapid iteration.
- Fragmented landscape, no clear winner.
- Monetisation is still evolving.
- Entry via niche use cases (tools, point solutions).

Lessons from Early Internet and CRO…

Case study #1: Early internet (互联网)

Initial phase: Free/low-cost user acquisition.

Monetisation follows scale (ads, e-commerce, platform take rates).

Winners built moats on:

Traffic and attention aggregation.

Proprietary user data.

Network effects and ecosystem lock-in.

Infrastructure layers (cloud, payments).

Key takeaway? Feature-level differentiation commoditised quickly; **data and ecosystem control drove pricing power**

Internet companies won on platform ecosystems; CROs on full-service integration

Case study #2: CRO evolution

Started with outsourced, low-value tasks.

Expanded into integrated, end-to-end R&D services.

Evolved into strategic, long-term partnerships.

Moats emerged from:

Integrated capabilities instead of single tools.

Regulatory expertise and execution credibility.

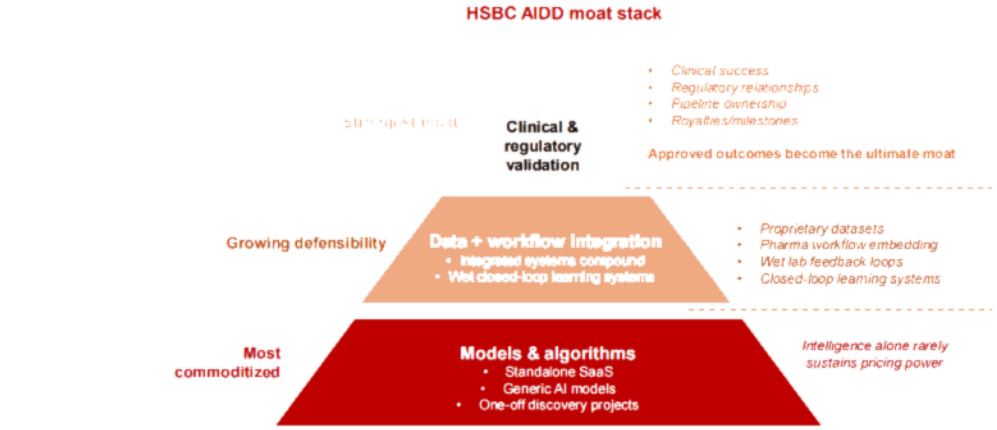
Scale and high switching costs.

Deep embedding in pharma workflows.

Key takeaway? Value shifted **from task execution to pharma workflow end-to-end integration.**

HSBC AI healthcare moat framework: Separating signal from noise

Exhibit 21. HSBC AI healthcare moat framework: Separating signal from noise



Source: Pharmcube, Company reports, HSBC Qianhai Securities

Exhibit 22. HSBC AI healthcare moat framework: Separating signal from noise

Likely to commoditise:	Likely to sustain pricing power:
◆ Standalone AI SaaS tools	◆ Companies controlling data generation loops
◆ Model-centric differentiation	◆ Deeply embedded players in pharma workflows
◆ One-off project-based revenue	◆ Platforms combining software + biology + automation
◆ Companies without proprietary data or wet lab integration	◆ Players with clinical and regulatory track record

Source: HSBC Qianhai Securities

Exhibit 23. HSBC AI healthcare moat framework: Which ones will last?

Moat	What it could look like?	Why it matters?
Durable moat: Data and closed-loop learning systems, feedback loops, not algorithms are compounding assets:	◆ Proprietary healthcare datasets (clinical, molecular, real-world evidence) ◆ AI models trained on exclusive data ◆ Feedback loops from: wet lab validation, clinical outcomes, iterative experimentation	◆ Creates self-reinforcing learning cycles ◆ Performance improves non-linearly over time ◆ Hard for competitors to replicate without equivalent data access
Monetisation moat: Platformisation, integrative system and ecosystem control	◆ From tool to workflow, platform and ecosystem ◆ Integration across the full pharma R&D process (discovery, clinical to manufacturing) ◆ AI, wet lab and automation (full-stack capability) ◆ API/data layer embedded across customers ◆ Multi-sided ecosystem (pharma, biotech, hospitals, payors)	◆ High switching costs ◆ Strong user stickiness ◆ Multiple monetisation levers (SaaS, services, downstream economics)
Weak moat: Models and algorithms	◆ Standalone AI SaaS tools ◆ Generalised, LLMs models	**Why do we see this as likely being commoditised with margin compression in standalone AI offerings?** ◆ Open-source model proliferation ◆ Fast benchmarking and replication ◆ Increasingly accessible compute ◆ Diffuse quickly and lose differentiation without proprietary data
Strongest moat: Clinical success and regulatory credibility	◆ Approved drugs ◆ Successful clinical trials ◆ Regulatory track record ◆ Pharma trust	◆ Converts AI capability into economic value ◆ Builds a credibility flywheel with pharma partners ◆ Enables long-term, strategic contracts (vs transactional revenue)

Source: HSBC Qianhai Securities

5. What is defensible and most valued in AI drug discovery, and what will likely be commoditised?

AIDD is a discipline that integrates machine learning (ML), deep learning (DL), and advanced statistical modelling to automate, optimise, and de-risk every stage of the preclinical and clinical drug development lifecycle. It represents a paradigm shift from legacy computer aided drug design (CADD) – which relies on rule-based computational chemistry and static molecular modelling – to a data-driven, iterative learning system that continuously improves predictive performance via closed-loop data integration.

Technically, AIDD has evolved through three peer-validated, distinct generations:

- First-generation (1970s-2000s): Rule-based quantitative structure-activity relationship (QSAR) models and linear regression frameworks, limited to single-parameter prediction of molecular properties with small, structured datasets.
- Second-generation (2000s-2019): Classical ML models – a machine learning algorithm that combines the predictions of multiple decision trees (random forests) to generate a single, more accurate result, support vector machines, gradient boosting) and shallow neural networks, enabling multi-parameter optimisation of molecular properties and virtual screening of large compound libraries.
- Third-generation (2020-present): Deep learning architectures, generative AI, multimodal large language models (MLLMs), and physics-informed neural networks, enabling end-to-end drug discovery from target identification to clinical trial design, with cross-modality data integration and closed-loop iterative learning.

Exhibit 24. Timeline of AIDD: From storytelling to data validation

Time period	Core technical breakthrough	Key players	Industry milestone
1970-80s	◆ First quantitative structure-activity relationship models ◆ 2D molecular descriptors	Tripos (founded 1979)	Birth of computer-aided drug design (CADD); first automated compound property prediction
1990s	◆ Molecular docking ◆ First-generation molecular force fields (AMBER, CHARMM) ◆ Structure-based drug design	Schrödinger (founded 1990), Biogen	First structure-based drug (Zanamivir, anti-influenza) approved; commercialisation of physics-based simulation tools
2000s	◆ High-throughput screening (HTS) ◆ Early machine learning for ADMET prediction ◆ First GPU-accelerated molecular simulation	Merck, Pfizer, Novartis, NVIDIA	Industrialisation of compound library screening; NVIDIA enters the space with GPU acceleration for computational chemistry, building the foundation for modern AIDD computing infrastructure
2010s	◆ Deep learning for molecular property prediction ◆ First generative molecular models ◆ AlphaFold 1 (first protein structure prediction breakthrough)	DeepMind (entered 2010s), Insilico Medicine (2014), XtalPi (2015)	Founding of the first modern AIDD start-ups; DeepMind enters the space with the first AlphaFold breakthrough; NVIDIA's GPU infrastructure becomes the standard for AIDD high-performance computing
2020-22	◆ AlphaFold 2 ◆ Diffusion models for molecular generation ◆ First AI-discovered candidates enter clinical trials	DeepMind, Isomorphic Labs, Exscientia, Insilico, XtalPi	DSP-1181 (first AI-designed molecule) enters Phase I; AlphaFold 2 solves the protein folding problem; first AIDD-discovered candidates advance to clinical trials
2023-24	◆ Closed-loop DMTA platforms ◆ Multi-objective generative AI ◆ First Phase II data for AI-discovered drugs	Insilico Medicine, XtalPi, Schrödinger	Insilico's rentosertib reports positive Phase IIa data; XtalPi's XFF force field validated with Pfizer; Schrödinger launches commercial FEP simulation tools
2025-26	◆ Open-source release of mature foundation models ◆ Multi-agent LLM systems ◆ Full open-source of AlphaFold 3 ◆ Open-source of BioNeMo	Google DeepMind, NVIDIA, OpenAI, AWS	AlphaFold 3 and NVIDIA BioNeMo are open-sourced, commoditising basic structure prediction; first agentic AI systems for autonomous R&D launch

Source: Company data, HSBC Qianhai Securities

To accurately value pure-play AIDD companies, investors must look beneat
of "AI". The prevailing market misconception is that AI merely accelerates tr
In reality, the true valuation premium is entirely dependent on the underlying
architecture – specifically, the transition from rigid-body screening to equiva
models and quantum-approximating neural networks.



Tier 1: The commoditisation of rigid-body docking (low defensibility)

The base layer of computational chemistry relies on structure-based virtual screening (SBVS). Historically, this involved connecting massive libraries (e.g., ZINC, Enamine REAL) against static protein structures. The open-sourcing of DeepMind's AlphaFold has effectively commoditised target structural data (source: DeepMind).

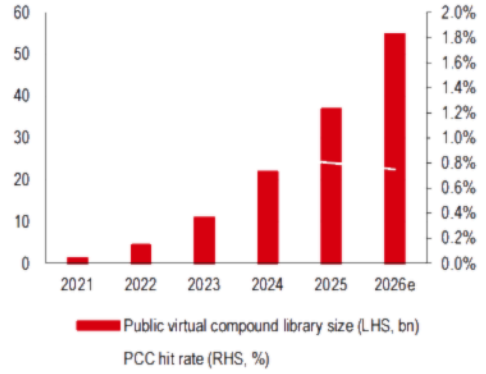
However, proteins are highly dynamic entities. Standard molecular docking algorithms rely on heuristic scoring functions that fundamentally fail to account for "induced-fit" mechanisms − the conformational changes a protein undergoes upon ligand binding. Companies that merely deploy basic convolutional neural networks (CNNs) to screen open-source structural databases are engaged in a low-margin computational arms race with Big Tech. Their predictive validity in complex human biology is inherently weak (source: Nature Computational Science).

Exhibit 25. Available protein structure has exceeded 200m

Year	Database	Number of known 3D protein structures
2020	PDB	c170,000
2021	AlphaFold v1	c350,000
2022-26	AlphaFold Protein Structure Database	>214,000,000

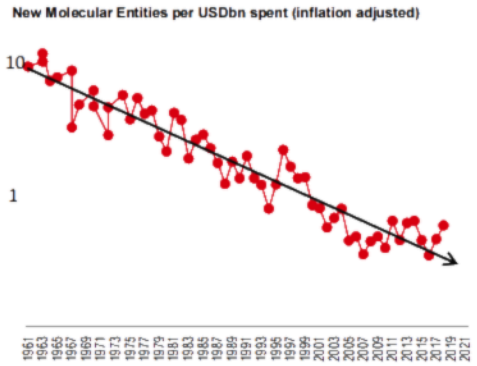
Source: AlphaFold Protein Structure Database, HSBC Qianhai Securities

Exhibit 26. Commoditisation of rigid-body docking are not helping with the preclinical candidate compounds (PCC) hit rate…



Source: ZINC20, Enamine Real, Reviews Drug Discovery, HSBC Qianhai Securities estimates

Exhibit 27. …per Eroom's law, drug discovery is becoming slower and more expensive



Source: Nature Reviews Drug Discovery, HSBC Qianhai Securities

Tier 2: Geometric deep learning and equivariant diffusion (the De Novo moat)

The first true algorithmic moat lies in generative architecture. Instead of screening existing libraries, advanced players use generative AI to synthesise entirely novel chemical matter.

The cutting-edge standard has shifted from simple generative adversarial networks (GANs) to equivariant graph neural networks (EGNNs) and 3D diffusion models. Equivariance is mathematically crucial: it ensures the neural network understands that a molecule's properties remain identical regardless of its 3D rotation or translation in space (source: ICLR).

Exhibit 28. Generative AI's compression of time to IND

R&D phase	Traditional pharma industry average (months)	AIDD enabled (months)
Target discovery to hit identification	c24	<2
Hit to lead optimisation	c30	c10
PCC to IND	c12	c6
Total time to IND	**c66**	**c18**

Source: Insilico, HSBC Qianhai Securities

Relevant players:

- **Insilico (3696 HK)**: Its proprietary Chemistry42 platform is a prime example of this moat. It uses over 30 generative engines, including structure-based diffusion models, to achieve multi-parameter optimisation (balancing affinity, ADME, and novelty simultaneously, source: Company reports).
- **Generate Biomedicines (GEMB US)**: Operates on a similar cutting-edge generative thesis but focuses on large molecules, using its Chroma platform – a generative diffusion model capable of programming programmable protein therapeutics based on structural constraints (source: Company reports).

Tier 3: Neural network potentials and hybrid physics-ML (the thermodynamic moat)

Once a molecule is generated, its binding affinity must be validated physically. Traditional quantum mechanics (e.g., density functional theory (DFT)) is highly accurate but computationally prohibitive at scale. Conversely, classical molecular dynamics (MD) is fast but lacks electronic-level accuracy.

The breakthrough moat here is the development of machine learning potentials (MLPs) or neural network potentials. These algorithms are trained on vast quantum mechanical datasets to approximate DFT-level accuracy at the computational cost of classical MD (source: Nature Chemistry).

Relevant players:

- **Schrödinger (SDGR US)**: The legacy pioneer of free energy perturbation (FEP+). Its moat remains defensible because its force fields (OPLS) have been refined over decades − a data barrier difficult for purely tech-native startups to cross.
- **XtalPi (2228 HK)**: Integrates advanced quantum physics calculations with an automated robotics wet-lab. XtalPi's moat is highly defensible because it creates a closed-loop "Tech-CXO" model; its algorithms predict crystalline structures, and its proprietary robotic arrays synthesise and validate them, generating high-fidelity, non-public data.
- **Iambic Therapeutics (US Private – acquired by Takeda in February 2026)**: Uses its NeuralPLexer technology, a state-of-the-art physics-informed neural network that predicts structural ensembles of protein-ligand complexes, capturing dynamic flexibility that standard AlphaFold predictions miss.

Tier 4: Phenomics and multimodal clinical prediction (The de-risking moat)

For equity investors, mitigating the Phase II to Phase III clinical failure rate (the efficacy proof-of-concept) dictates valuation. Biological complexity cannot be solved by chemical structures alone. The ultimate moat involves mapping the drug to human systems biology.

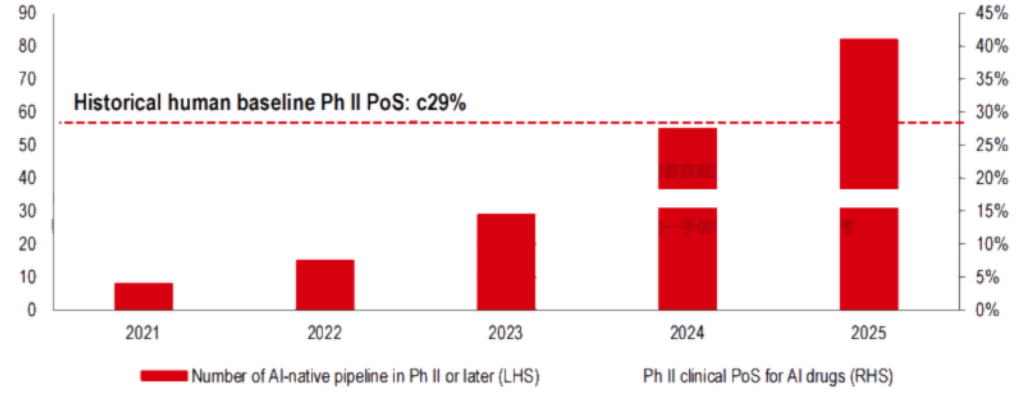
Recursion Pharmaceuticals (RXRX US) pioneers the phenomics approach. Instead of focusing solely on target binding, Recursion uses automated "cell painting" to capture high-dimensional morphological changes in cells upon drug exposure. Their deep learning models map a massive biological dataset (over 50 petabytes) to identify novel biological relationships (source: Cell Systems).

Exhibit 29. AI vs human clinical success rates (PoS premium)

Clinical metrics	Historical human R&D baseline (PoS)	AI-assisted clinical trial success rate
Ph I success rate (safety)	40-65%	80-90%
Ph II success rate (efficacy)	c28.9%	c40%
End to end probability of success	5-10%	9-18%

Source: BCG, Biotechnology Innovation organisation, Informa Pharma Intelligence, HSBC Qianhai Securities

Exhibit 30. Multimodal de-risking and PoS premium



Source: ClinicalTrials, Informa Pharma Intelligence, BCG, HSBC Qianhai Securities

Relevant players:

- **Insilico (3696 HK)**: Its inClinico engine exemplifies multimodal clinical predictive AI. It leverages multi-omics data, natural language processing (NLP) of global clinical trials, and proprietary knowledge graphs (KGs) to forecast Phase II/III success rates, offering a quantifiable probability of success (PoS) premium for DCF valuations.
- **Tempus (TEM US)**: Focuses heavily on real-world evidence (RWE). Its moat relies on amassing the world's largest library of clinical and molecular data, specifically in oncology, applying AI to discover novel biomarkers and optimise clinical trial patient stratification.

Exhibit 31. Comparable table

Company	Ticker	Market cap (USDm)	T/O (USDm)	Currency	Price (LC)	Rating	TP (LC)	2026e PS (x)	2027e PS (x)	2026e PB (x)	Revenue growth 2027e	Revenue growth 2028e	L5Y avg PS (x)	L5Y min PS (x)	L5Y max PS (x)	2027e PSG
AI SaaS/platform and service focus																
XtalPi	2228 HK	3,870	64	HKD	7.05	Hold	8.00	21.5	15.6	2.8	38	34	33.4	16.6	59.5	0.6
Schrödinger	SDGR US	1,128	18	USD	15.10			4.8	4.5	3.7	8	25	14.3	3.2	60.4	0.6
Certara	CERT US	1,066	24	USD	6.85			2.7	2.6	1.1	2	7	8.6	1.7	23.8	1.5
Tempus AI	TEM US	8,795	306	USD	48.98			5.5	4.5	23.1	23	20	8.4	2.8	16.1	0.2
Average								**8.6**	**6.8**	**7.7**	**18**	**21**	**16.2**	**6.1**	**39.9**	**0.7**
Pipeline centric (AI biotechs)																
Insilico Medicine	3696 HK	3,266	44	HKD	44.24	Buy	84.50	15.4	11.3	1.0	34	19	27.9	15.1	37.5	0.5
Recursion Pharmaceuticals	RXRX US	1,632	60	USD	3.08			24.7	14.5	2.1	71	89	55.0	17.6	201.2	0.3
Absci	ABSI US	1,405	36	USD	8.32			NA	78.0	6.4	20	76	90.7	8.6	894.0	NA
Intellia Therapeutics	NTLA US	1,555	81	USD	11.47			27.2	17.2	2.4	58	130	68.1	10.3	348.2	0.5
AbCellera Biologics	ABCL US	1,823	35	USD	5.97			57.1	58.9	NA	-3	8	22.5	3.3	102.8	NA
CRISPR Therapeutics	CRSP US	4,644	90	USD	48.15			188.6	33.9	2.5	456	276	727.5	1.8	11216.9	0.4
Average (excl. CRISPR)								**31.1**	**36.0**	**2.9**	**36**	**64**	**52.8**	**11.0**	**316.8**	**0.4**

Note: Priced as of 21 July 2026.
Source: Bloomberg consensus estimates, HSBC Qianhai Securities estimates

Valuation and risks

		Valuation	Risks
Insilico Medicine 3696 HK **Buy**	Current price: HKD44.24 Target price: HKD84.50 Upside: +91.0%	We continue to use a SOTP model to value Insilico. Our key assumptions include a WACC of 9.6% (based on a risk-free rate of 4.25% and a market risk premium of 4.75%), a terminal growth rate of 3.0%, and an end-2026e USD/HKD exchange rate assumption of 7.83 (all unchanged). Our TP of HKD84.50 implies 91% upside from current levels; accordingly, we maintain our Buy rating on the stock.	**Key downside risks** ◆ Clinical trial failure in key asset. ◆ Delays in platform output or weak deal-making activity. ◆ Unpredictable revenue and milestone timing. ◆ Technological disruption. ◆ Intense competition. ◆ Partner dependency. ◆ Revenue concentration. ◆ Geopolitical risk.
		Linda Shu*, PhD (Reg. No. S1700522120001) | linda.y.l.shu@hsbcqh.com.cn | +86 755 88983246	
XtalPi 2228 HK **Hold**	Current price: HKD7.05 Target price: HKD8.00 Upside: +13.5%	We continue to use a DCF model to value XtalPi. Our key assumptions include a WACC of 9.9% (based on a risk-free rate of 4.25% and a market risk premium of 4.75%), a terminal growth rate of 3.0% and an end-2026e RMB-HKD exchange rate assumption of 1.17 (all unchanged). Our TP of HKD8.00 implies c13.5% upside from current levels. We maintain our Hold rating on the stock because we view its current valuation as fair, which reflects both near-term opportunities and execution risks.	**Key upside risks** ◆ Faster-than-expected major deals' milestone achieved. ◆ Major new MNC partnership. ◆ Robotics capacity expansion completion. ◆ Pipeline company IPO/M&A. ◆ Operating leverage inflection. **Key downside risks** ◆ Milestone payment delay. ◆ Competition intensifies. ◆ Regulatory risk. ◆ Forex risk.
		Linda Shu*, PhD (Reg. No. S1700522120001) | linda.y.l.shu@hsbcqh.com.cn | +86 755 88983246	

Note: Priced as of 21 July 2026. **Employed by a non-US affiliate of HSBC Securities (USA) Inc, and is not registered/ qualified pursuant to FINRA regulations.*
Source: Wind, HSBC Qianhai Securities estimates

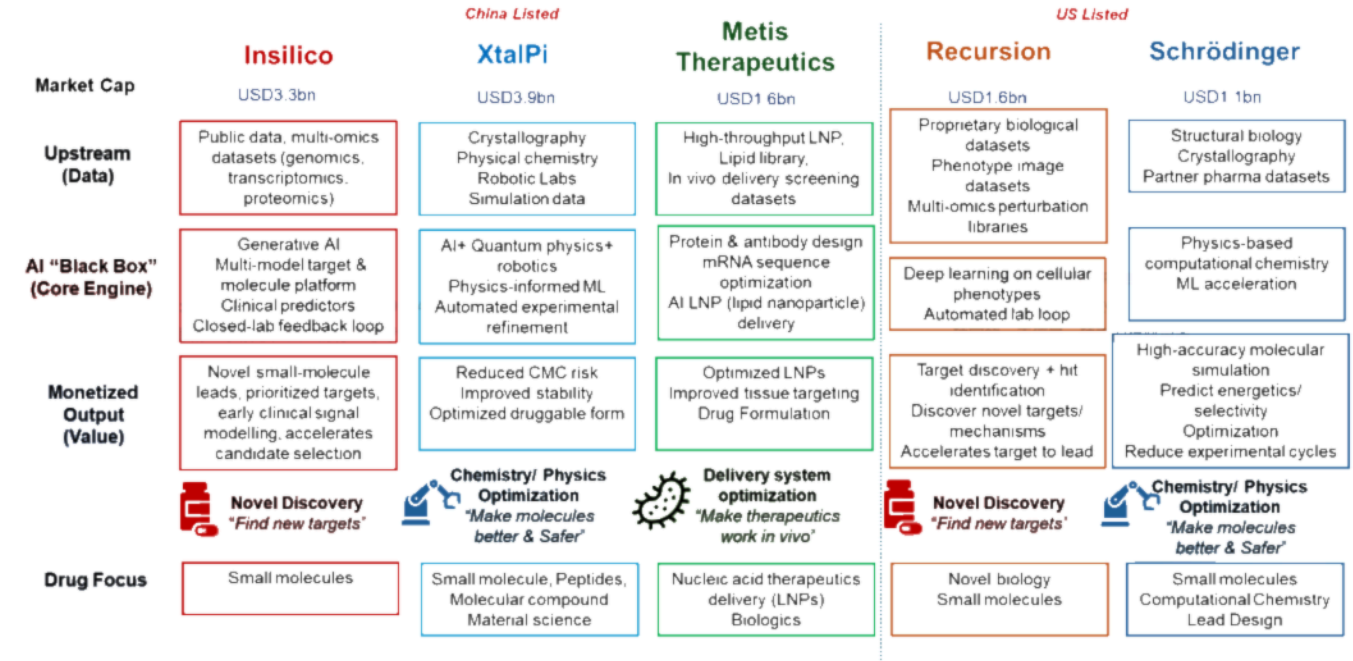
Appendix

Exhibit 32. Snapshot of key global players in AI + healthcare

Company	Business model	Proprietary dataset?	Wet lab integration?	Proven BD / MNC partnerships?	Owns clinical assets?
SaaS/Platform and service focus					
XtalPi	AI + quantum simulation + robotics platform monetising automated chemistry/CRO workflows and autonomous lab operations.	Yes	Yes	Yes	Limited
Schrödinger	Physics-based molecular modeling SaaS combined with partnered/internal drug discovery programmes.	Yes	Partial	Yes	Yes
Certara	Regulatory-focused biosimulation SaaS platform enabling model-informed drug development and FDA submissions.	Yes	No	Yes	No
Tempus AI	Precision medicine platform monetising genomic diagnostics, multimodal patient data, and AI-driven clinical insights.	Yes	Yes	Yes	Limited
Pipeline centric (AI biotechs)					
Insilico Medicine	Generative AI drug discovery company designing novel targets and molecules for partnered and internal pipelines.	Mostly Public	Yes	Yes	Yes
Recursion Pharmaceuticals	Industrialised AI-biotech platform using automated wet labs to generate proprietary biological datasets for drug discovery.	Yes	Yes	Yes	Yes
Absci	AI biologics design company combining generative AI with synthetic biology to engineer therapeutic proteins.	Yes	Yes	Yes	Yes
Intellia Therapeutics	CRISPR gene-editing biotech transitioning from platform/licensing toward fully owned therapeutic pipeline.	Yes	Yes	Yes	Yes
AbCellera Biologics	AI-enabled antibody discovery platform evolving from fee-for-service discovery into pipeline ownership.	Yes	Yes	Yes	Yes
CRISPR Therapeutics	Gene-editing therapeutics company commercialising CRISPR/Cas9-based medicines through internal pipeline development.	Yes	Yes	Yes	Yes
Exscientia	AI-driven precision drug design platform focused on computational molecule optimisation and partnered pipelines.	Yes	Partial	Yes	Yes
DeepMind	Foundational AI research platform whose AlphaFold models transformed protein structure prediction infrastructure.	Mostly Public	No	Yes	No
Insitro	Machine-learning-first biotech integrating proprietary human datasets with wet-lab validation for internal therapeutics.	Yes	Yes	Yes	Yes
BenevolentAI	AI-enabled target discovery platform combining knowledge graphs with internal biological validation capabilities.	Yes	Yes	Yes	Yes

Source: Company data, HSBC Qianhai Securities

Exhibit 33. Technological comparison of major players in AI + healthcare



Source: Company data, HSBC Qianhai Securities