

Reducing Early Portfolio Risk with Retrosynthesis

How AI-enabled synthetic intelligence can inform strategic program decisions

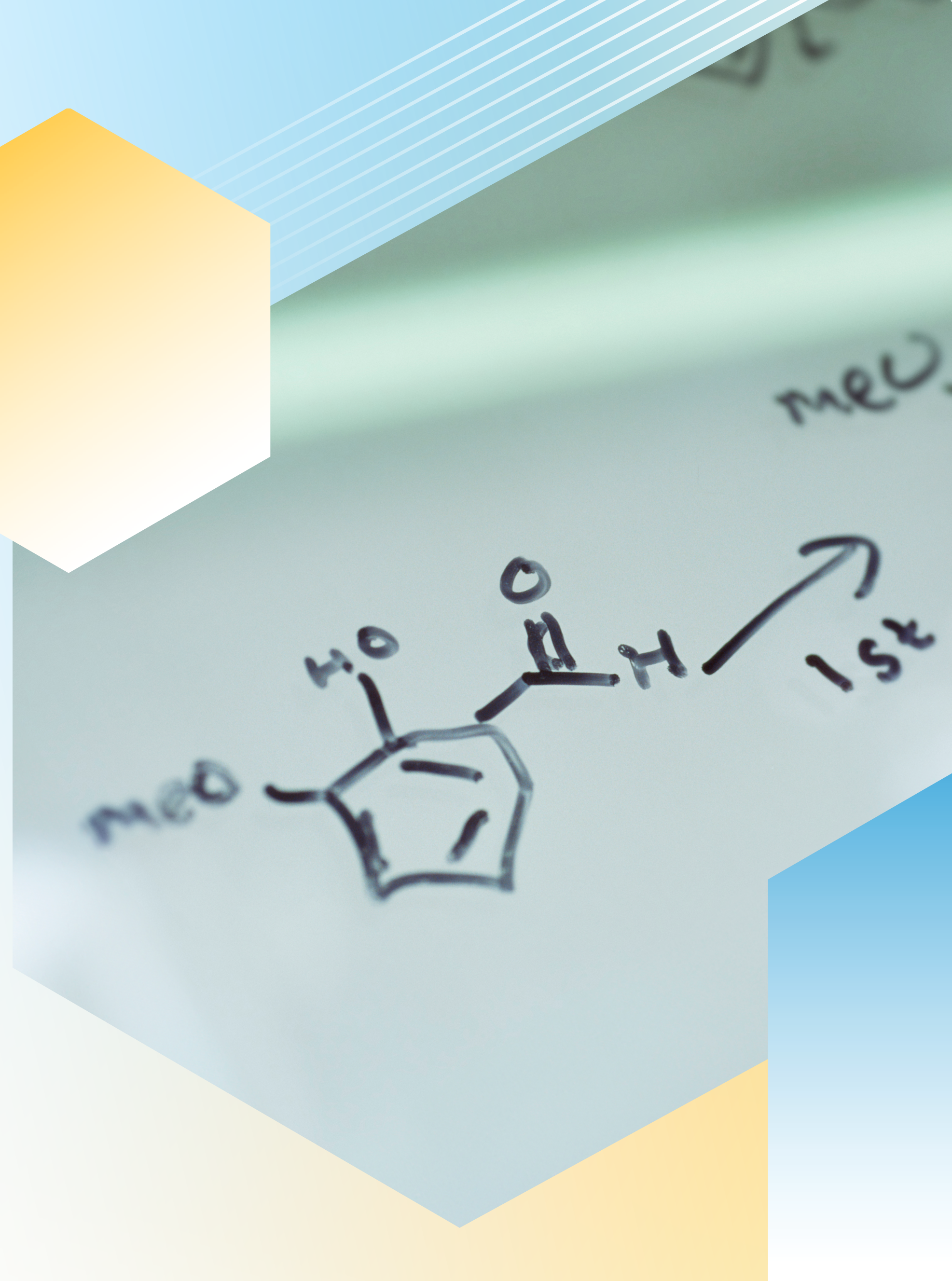


Table of Contents

3

Executive summary

4

The escalating pressure on early portfolio decisions

5

Synthetic feasibility as a source of portfolio risk

7

Retrosynthesis: a tool for feasibility assessment

9

Retrosynthesis tools: looking beyond early deployment

10

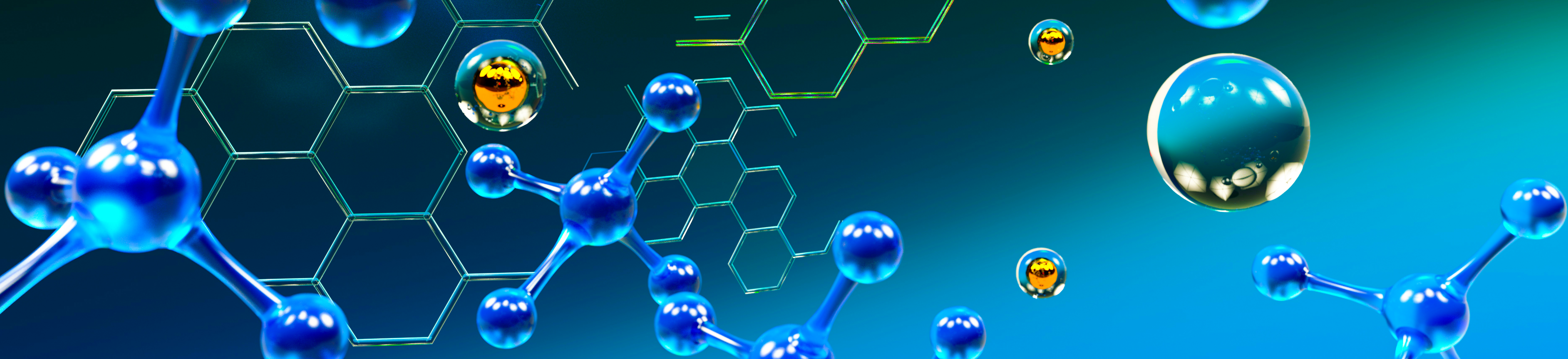
Strengthening early portfolio decisions
with integrated feasibility insight

11

References

Executive summary

Rising costs, low success rates, and increasing competition are intensifying pressure on early drug development decisions. Yet synthetic feasibility is often overlooked as a source of early portfolio risk, leading to costly delays or failures later. This whitepaper highlights feasibility as a key source of portfolio risk and shows how early retrosynthetic analysis can improve candidate selection by revealing route complexity, precedent, and supply constraints. It also explores how integrating AI-enabled retrosynthesis into early workflows can help organizations make more informed, lower-risk decisions, thereby improving capital efficiency and development success.



The escalating pressure on early portfolio decisions

Pharmaceutical development costs are drastically increasing. With the average R&D investment per asset now estimated at US\$2.23 billion, every advancement decision carries a substantial financial risk. At the same time, indications are becoming increasingly crowded.¹ The average number of assets per target rose from 3 in 2000 to 7 in 2022, a jump compounded in already competitive indications such as oncology, which climbed from 1.8 to 9 over the same period.²

In this environment, late-stage failures and reprioritizations are increasingly devastating, due to larger financial losses and more competitors poised to capitalize on any delays or failures. As such, development teams are dealing with intensifying capital-efficiency expectations from boards and investors, as well as growing scrutiny of portfolio governance—yet attrition rates remain high. Indeed, one recent longitudinal analysis of 20,398 clinical development programs reported clinical trial success rates ranging from 7% to 20%, highlighting the critical need for developers to take steps to increase the likelihood of program success.³

One driver of these attrition rates that receives insufficient attention is synthetic feasibility. Without early visibility into synthetic complexity, even when compounds show promising biological profiles, downstream difficulty in synthesis can prevent optimization and scale-up, leading to program abandonment or deprioritization.⁴ This challenge will only be amplified as AI-driven expansion of chemical space continues to spread across the pharma industry, as generated compounds are widely known to be difficult or impossible to synthesize.⁵

This whitepaper will discuss how synthetic feasibility can be a key source of program risk, highlight how early portfolio decisions can benefit from retrosynthetic insights, and recommend actionable approaches to successfully incorporate AI-enabled retrosynthetic tools into pharma workflows.

Synthetic feasibility as a source of portfolio risk

Early portfolio decisions typically prioritize biological potency, selectivity, target engagement, and safety signals. However, the “make” step is often the most time-consuming and costly stage of the Design-Make-Test-Analyze (DTMA) cycle, with synthesis accounting for about 60% of the total DMTA cycle time (Table 1).^{6,7}

Table 1: Typical DMTA cycle timeline for small molecule lead optimization

Cycle Stage	Activities	Typical Time Taken
Design	Medicinal chemistry ideation and computational scoring	2 days
Make	Synthesis, purification, and quality control	21 days
Test	Biochemical, cellular, and Tier 1 ADME panel tests	10 days
Learn	Data analysis and team review	2 days
Total		35 days/ ~5 weeks

To enable successful commercialization, assets must have a viable synthesis route that takes into account:

- Route complexity, including step count and functional group compatibility
- Reaction precedent and robustness
- Availability and cost of starting materials
- Scalability of intermediates
- Route flexibility and optionality
- Sensitivity to supply chain disruptions
- Downstream cost-of-goods implications

Boots vs. BHC: The impact of synthetic route decisions

Synthetic routes have far-reaching consequences beyond feasibility, affecting cost, yield, and environmental impact. The shift from the six-step Boots process (~40% atom economy) to the three-step BHC process (~77% atom economy) for ibuprofen production illustrates this clearly. The BHC process nearly doubles yield, generates recoverable byproducts rather than waste, and uses recyclable catalysts, creating a material impact at an industrial scale.⁸

Synthetic routes are most commonly explored during lead optimization, with little to no synthetic insight included in early decision-making.⁴ This approach means that unanticipated yield instability, scale-up failures, impurity challenges, and reliance on scarce or unstable intermediates do not surface until lead optimization. Ultimately, these challenges can result in route redesign or even candidate abandonment, eroding investor confidence through sunk costs and delays.

Additionally, insights are generally left tacit, distributed across chemists’ experience, literature searches, patents, internal ELNs, and commercial catalogs, making it difficult to compare synthetic factors across candidates to fuel better decisions.

By connecting synthetic insight to executive decision drivers and taking a structured approach to comparing synthetic feasibility across candidates, organizations can begin to close this knowledge gap and mitigate portfolio risk.



Comparing synthetic feasibility across candidate scaffolds

Beyond risk mitigation, synthesis feasibility can also offer another selection criterion. When multiple chemotypes demonstrate promising biological activity, synthetic feasibility assessment provides teams with an additional way to compare candidates beyond biological promise. Table 2 illustrates a simplified comparison between two hypothetical scaffolds targeting the same biological pathway.

Table 2: Example candidate scaffold feasibility comparison.

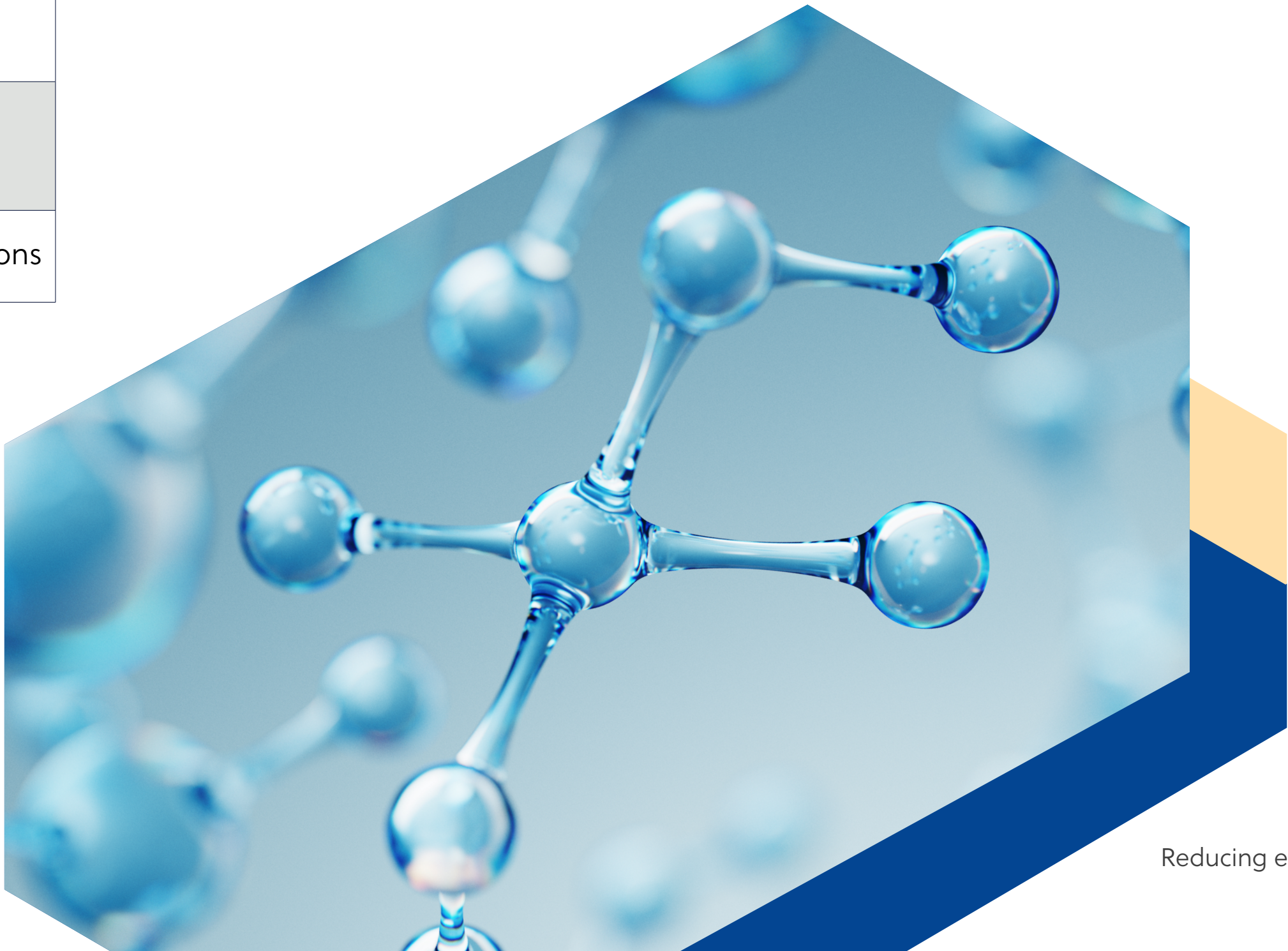
Feasibility Dimension	Scaffold A	Scaffold B
Route Complexity	9 steps with multiple protecting groups	5-step convergent route
Reaction Precedent	Limited literature precedent for key cyclization	Well-established transformations
Starting Materials	Two intermediates require custom synthesis	Commercially available starting materials
Route Optionality	Single route family identified	Multiple viable disconnections

Viewed individually, each of these factors may appear manageable. However, taken together, they reveal meaningful differences in development risk.

Scaffold A concentrates risk in several areas simultaneously: longer route length, limited precedent for a key transformation, and reliance on custom intermediates. This combination increases the likelihood that route redesign or supply constraints could emerge later in development.

Scaffold B distributes risk more favorably: A shorter convergent route supported by well-established reactions reduces experimental uncertainty, while commercially available starting materials and multiple disconnection strategies provide greater flexibility if challenges arise.

For portfolio teams, synthetic feasibility comparisons, such as those enabled by retrosynthesis, can identify the most robust development pathways, thereby strengthening prioritization decisions.

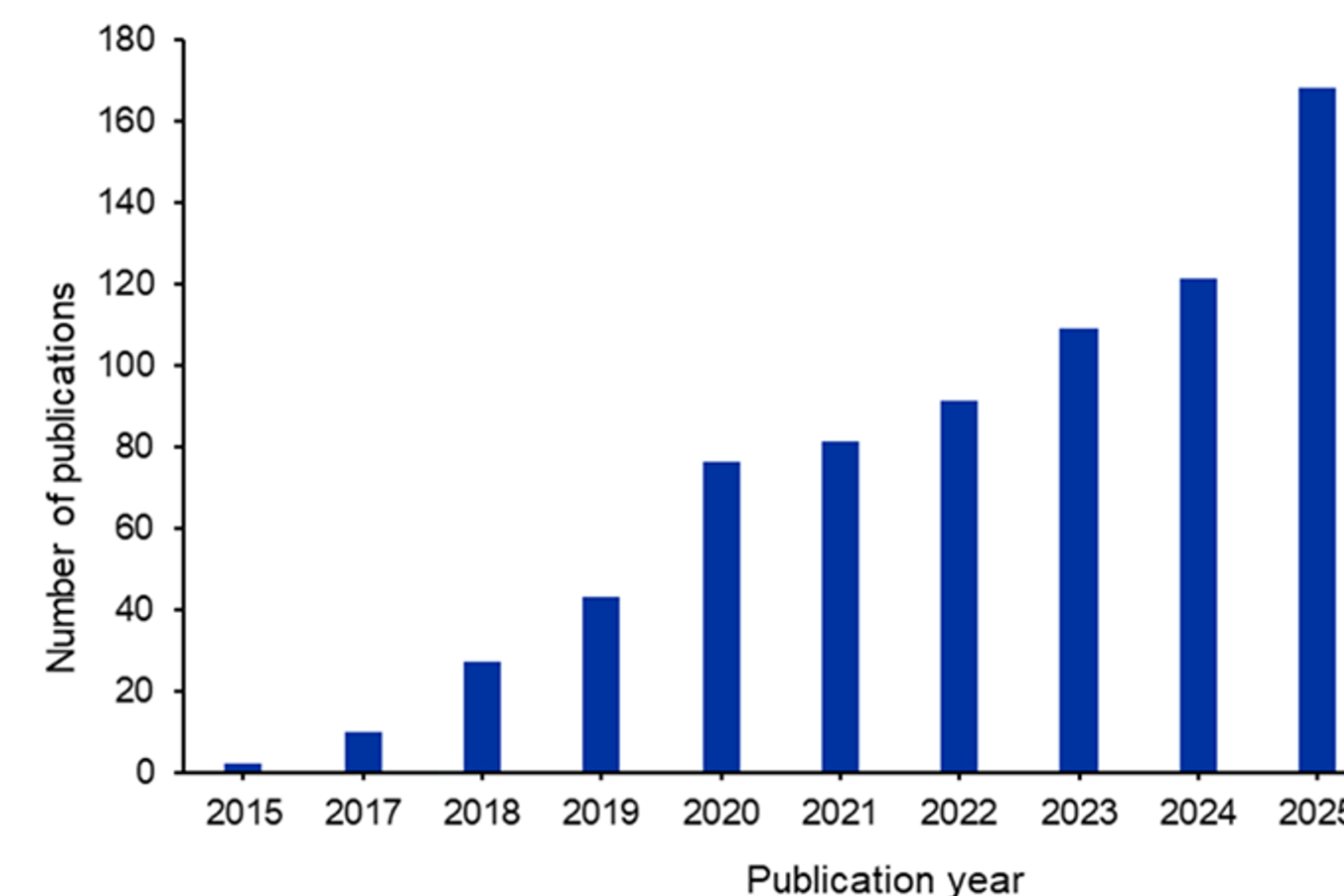


Retrosynthesis: a tool for feasibility assessment

Retrosynthesis is a powerful, foundational approach for assessing synthetic feasibility, translating a target molecule into a set of choices. By systematically deconstructing the molecule into simpler precursors, retrosynthesis proposes multiple disconnection strategies, maps reaction precedents, evaluates route optionality, and helps identify potential bottlenecks early. When used well, it can reveal whether a scaffold depends on unusual starting materials, whether a key transformation rests on thin precedent, whether a route family has credible backups, and where time-to-materials may become the hidden constraint.

Growing interest and research into AI-enabled retrosynthesis has also raised the bar for what is possible (Figure 1). Route planning has been improved with strategies including disconnection prompting, which can be used to steer models towards more diverse and useful disconnections, delivering a 39% performance improvement; molecular string editing to improve route diversity while maintaining accuracy; and decision transformer models for multistep route planning.⁹⁻¹¹ Reaction-level prediction is also advancing, with yield prediction used to reduce waste and deliver more relevant compounds through the design-make-test-analyze cycle.¹²

Figure 1: Yearly trends of patent families and journal articles published on AI in retrosynthesis-related research. Data includes journal and patent publications from the [CAS Content Collection](#) for the period 2016-2025.



AI in retrosynthesis is most effectively used at stages requiring rapid exploration and decision support, such as early-stage molecule design, multi-step route planning, pathway exploration, and feasibility assessment before experimental work.¹³ In practice, AI systems can predict reaction steps and reagents, generate multiple candidate synthetic routes, and evaluate and rank pathways based on feasibility, yield, and cost, enabling efficient identification of practical strategies.¹⁴ However, expert chemist validation and oversight are essential when using AI retrosynthesis tools, as they often produce numerous possible pathways and may lack complete chemical intuition.

“AI in retrosynthesis should be used throughout the workflow, from early molecule screening and route generation to optimization and feasibility assessment, where it accelerates decision-making, explores multiple pathways, and reduces synthetic risk.”

[Gary Gustafson](#), PhD. Senior Customer Success Specialist, CAS.

In many organizations, retrosynthesis use is confined to tactical route planning, lead optimization support, problem-solving when synthesis stalls, and bench-level execution refinement. Additionally, retrosynthesis insights are rarely standardized across candidates at portfolio checkpoints, structured into comparable feasibility outputs, or translated into portfolio risk language, making them ineffective for supporting confident early portfolio decisions and mitigating program risk.

To realize the true value of retrosynthesis for portfolio decisions, it must be integrated early into workflows, when scaffold selection and program optionality remain high. At these early stages, retrosynthesis outputs can be translated into structured indicators such as route complexity, reaction precedent depth, starting-material accessibility, and route optionality (Table 3).

Table 3: Retrosynthesis applications for early-stage portfolio decisions.

Early Discovery Stage	Key Retrosynthesis Questions	Strategic Value
Hit-To-Lead Exploration	Are proposed chemotypes supported by viable synthetic routes?	Eliminates synthetically fragile scaffolds early
Series Prioritization	Do competing scaffolds differ significantly in route complexity or precedent depth?	Enables chemistry-informed prioritization
Pre-Portfolio Review	Are viable route families available for the most promising candidates?	Surfaces feasibility risks before program commitment

These indicators enable organizations to identify hidden feasibility risks, thereby supporting balanced portfolio decisions and resilient development pathways.



Retrosynthesis tools: looking beyond early deployment

Early deployment alone is not sufficient to support portfolio decisions; teams also need to select the right retrosynthesis tools. Ranging from simple retrosynthesis predictions to integrated AI-enabled retrosynthesis platforms, teams must ensure the retrosynthesis tool they use has several core capabilities to optimally inform portfolio governance.

Reliable evidence-based prediction

Reliable decisions require reliable evidence-based prediction. Therefore, route generation must be grounded in both quality and scale. In practice, a large volume of diverse reaction data and underlying catalogued substances are needed to support both experimental and predictive retrosynthesis pathways and avoid narrow precedent bias. Each proposed transformation must also be grounded in real-life chemistry expertise, ensuring it is traceable to experimental precedent and the full range of experimental context, including reagents, catalysts, conditions, yields, and stereochemical outcomes.

Chemistry-aware interactivity

To enable the flexibility needed for early portfolio decisions, retrosynthesis tools require interactive filtering tailored to route optimization. Filters should include more common options, such as cost, step count, and feasibility, as well as complex tools for filtering based on stereochemical and functional-group considerations, starting-material availability, and the exclusion of undesirable reagents or reaction classes.

Structured comparison enablement

Beyond filtering, interactive retrosynthesis tools must also enable structured comparison across candidates to standardize and streamline synthetic feasibility portfolio decisions. Tools should enable comparison of route length and complexity, depth of literature precedent, predicted feasibility or confidence, starting material availability and supply sensitivity, cost exposure, and scalability risk.

Broad synthetic intelligence integration

Retrosynthesis alone can enumerate options, but portfolio decisions require a broader context. As such, retrosynthesis tools must integrate data from scientific journals, patent information, manufacturer details, global regulations, and environmental considerations to enable more streamlined, repeatable feasibility decisions. However, many retrosynthesis tools lack the comprehensiveness needed to support an integrated approach, leaving synthetic intelligence scattered across sources. This fragmentation creates avoidable blind spots and limits an organization's ability to review precedent comprehensively, cross-compare candidates, and confidently assess route optionality.

Strengthening early portfolio decisions with integrated feasibility insight

As competitive pressure and development costs increase, early portfolio decisions increasingly define downstream exposure to cost, time, and supply risk. Though widely used, retrosynthesis has under-realized strategic value for early portfolio decisions. When applied deliberately at early decision points and supported by integrated synthetic intelligence, retrosynthesis enhances transparency, comparability, and alignment. As such, organizations that leverage advanced retrosynthesis tools and integrated synthetic intelligence into early decision points are better equipped to evaluate feasibility risk before it compounds, enabling lower risk and costs and ultimately improving competitive standing.

Discover AI-driven retrosynthesis with the CAS SciFinder Discovery Platform™

The CAS SciFinder Discovery Platform™ enables chemists to apply AI-driven retrosynthesis to explore potential synthetic routes, evaluate reaction precedent, and assess feasibility earlier in the discovery process. Powered by the CAS Content Collection™, it connects retrosynthetic planning with trusted data from journals, patents, and suppliers to support more informed R&D decisions. CAS is a division of the American Chemical Society.

[Explore more](#)



References

1. Navigating the GLP-1 boom | Deloitte. Accessed June 5, 2026. <https://www.deloitte.com/us/en/industries/life-sciences-health-care/perspectives/navigating-the-glp-boom.html>
2. Herding in the drug development pipeline. Accessed June 5, 2026. <https://www.nature.com/articles/d41573-023-00063-3>
3. Zhou Y, Zhang Y, Xu H, et al. Dynamic clinical trial success rates for drugs in the 21st century. Nat Commun. 2025;16(1):9537. doi:10.1038/s41467-025-64552-2
4. Advancing Cancer Drug Discovery: Bypassing Synthesis Bottlenecks - European Society of Medicine. March 1, 2026. Accessed June 5, 2026. <https://esmed.org/advancing-cancer-drug-discovery-bypassing-synthesis-bottlenecks/>
5. Chen S, Nam G, Aspuru-Guzik A, Jung Y. SynTwins: a retrosynthesis-guided framework for synthesizable molecular analog generation. Chem Sci. 2026;17(4):2255-2262. doi:10.1039/D5SC05225D
6. Synthesis AI Part 1: The Synthesis Bottleneck — Why “Make” Lags Behind. Seongok Ryu. April 4, 2026. Accessed June 5, 2026. <https://seongokryu.github.io/posts/synthesis-ai-part1-synthesis-bottleneck/>
7. Nippa DF, Boddy AJ, Atz K, Grether U, Binch H, Martin RE. Accelerating compound synthesis in drug discovery: the role of digitalisation and automation. RSC Med Chem. 2025;16(12):5753-5764. doi:10.1039/D5MD00672D
8. The Synthesis of Ibuprofen: A Technical Guide to the Boots and BHC Processes. Accessed June 5, 2026. https://pdf.benchchem.com/25/The_Synthesis_of_Ibuprofen_A_Technical_Guide_to_the_Boots_and_BHC_Processes.pdf
9. Granqvist E, Mercado R, Genheden S. Retrosynformer: planning multi-step chemical synthesis routes via a decision transformer. Digit Discov. 2026;5(1):348-362. doi:10.1039/D5DD00153F
10. Thakkar A, Vaucher AC, Byekwaso A, Schwaller P, Toniato A, Laino T. Unbiasing Retrosynthesis Language Models with Disconnection Prompts. ACS Cent Sci. 2023;9(7):1488-1498. doi:10.1021/acscentsci.3c00372
11. Han Y, Xu X, Hsieh CY, et al. Retrosynthesis prediction with an iterative string editing model. Nat Commun. 2024;15(1):6404. doi:10.1038/s41467-024-50617-1
12. Incorporating Synthetic Accessibility in Drug Design: Predicting Reaction Yields of Suzuki Cross-Couplings by Leveraging AbbVie's 15-Year Parallel Library Data Set | Journal of the American Chemical Society. Accessed June 5, 2026. <https://pubs.acs.org/doi/10.1021/jacs.4c00098>
13. Mo Y, Guan Y, Verma P, et al. Evaluating and clustering retrosynthesis pathways with learned strategy. Chem Sci. 2021;12(4):1469-1478. doi:10.1039/D0SC05078D
14. A comprehensive survey of AI-based retrosynthesis planning: Datasets, models, and tools. Accessed June 5, 2026. <https://www.the-innovation.org/article/doi/10.59717/j.xinn-inform.2026.100026>

CAS connects the world's scientific knowledge to accelerate breakthroughs that improve lives. We empower global innovators to efficiently navigate today's complex data landscape and make confident decisions in each phase of the innovation journey. As a specialist in scientific knowledge management, our team builds the largest authoritative collection of human-curated scientific data in the world and provides essential information solutions, services, and expertise. Scientists, patent professionals, and business leaders across industries rely on CAS to help them uncover opportunities, mitigate risks, and unlock shared knowledge so they can get from inspiration to innovation faster. CAS is a division of the American Chemical Society.

Connect with us at cas.org