

From Evidence to Decisions: How AI Can Help Pharma Reach Better Development Decisions Earlier

A practical framework for earlier program-development decisions on when to stop, pivot, learn - or accelerate

Rafael Ioschpe | August 2026

In an earlier article, I argued that artificial intelligence could progressively transform Pharma into an evidence industry - one in which competitive advantage increasingly depends on the ability to generate, connect, interpret and validate evidence faster.

That transformation is already under way. AI is being applied to quantitative pharmacology, Model-Informed Drug Development (MIDD), disease-progression modeling, clinical-trial prediction and multimodal evidence integration. Recent scientific literature increasingly describes AI and MIDD as complementary rather than competing approaches. [1-4]

But there is a harder question behind the idea of Pharma becoming an evidence industry:

What happens after the evidence is generated?

More data do not automatically produce better decisions. More frequent interim analyses do not necessarily identify emerging problems earlier if the information being collected remains insufficiently discriminative. An AI model capable of processing biologically ambiguous evidence faster may therefore accelerate analysis without materially reducing the uncertainty surrounding the development decision.

An evidence industry without a decision-architecture risks becoming a data industry.

The opportunity explored here is not primarily to build an AI capable of deciding which drug should live or die. It is to design clinical development so that evidence capable of changing a decision becomes available earlier - and then determine where AI can help find, integrate, interpret and challenge that evidence without obscuring scientific accountability.

Two different problems: Evidence Latency and Decision Latency

When emerging evidence begins to reduce confidence in a development program, familiar responses include additional interim analyses, adaptive designs, futility criteria and more frequent data reviews.

All can be valuable.

But consider a program in which the clinically discriminative readout appears only after 18 months. Reviewing the trial at months 6, 9 and 12 may simply provide three earlier views of information that is still incapable of answering the critical development question.

In this article, Evidence Latency is defined as:

the time between a critical development hypothesis and the point at which sufficiently informative evidence becomes available to materially change a development decision.

A second bottleneck, Decision Latency, is defined as:

the time between sufficiently informative evidence becoming available and an accountable development action.

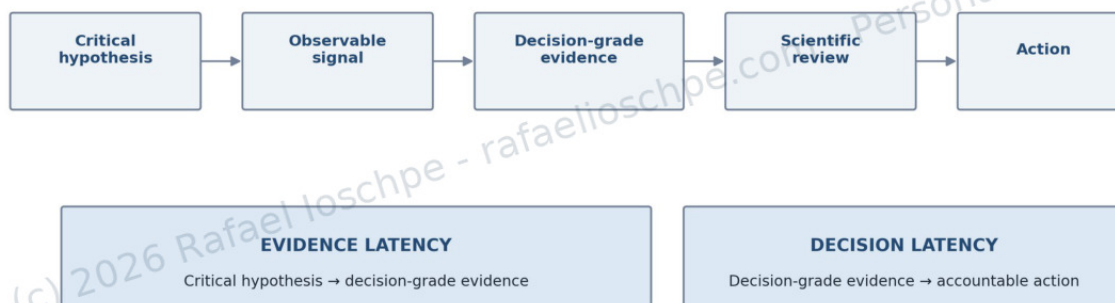
These are descriptive constructs rather than standardized regulatory terms. Their value lies in separating two problems that require different solutions.

Evidence Latency is predominantly an information problem. Biomarkers, longitudinal measurements, PK/PD, disease models, MIDD, Bayesian inference and AI may help shorten it.

Decision Latency is predominantly a decision-design and governance problem. Predefined decision boundaries, explicit ownership, structured overrides, independent challenge and mandatory reassessment become more relevant.

AI may help compress Evidence Latency. It cannot, by itself, resolve Decision Latency.

Figure 1. Drug Development Has Two Different Latency Problems



AI can help compress Evidence Latency. Governance is required to compress Decision Latency.

Three development cases, three different decision problems

The distinction becomes clearer through three real development examples. They were not selected as "AI success stories." Their value is precisely that they show how quantitative evidence and development architecture can alter decisions even before AI is introduced.

Pfizer / domagrozumab - when early evidence and later investment decisions diverge

Domagrozumab was an anti-myostatin monoclonal antibody developed for Duchenne muscular dystrophy. Pfizer prospectively evaluated an anabolic proof-of-mechanism signal through change in lean body mass using a Bayesian decision framework.

The early study demonstrated pharmacological activity, but the evidence did not reach the level prospectively defined as sufficient for success. Pfizer nevertheless proceeded to a larger Phase II proof-of-concept study. In the company's subsequent analysis of R&D productivity, the authors described competitive considerations and the importance of maintaining development momentum as factors contributing to the decision. The Phase II study subsequently failed to demonstrate the anticipated clinical efficacy, and development was discontinued. The authors later suggested that a smaller Early Signal of Efficacy study might have addressed the relevant development question earlier. [5]

This should not be interpreted as evidence that a quantitative criterion should mechanically determine the future of an asset. Development decisions legitimately integrate scientific, clinical, competitive and strategic information that may extend beyond an individual model.

The broader lesson is more subtle. When emerging evidence diverges from prospectively established expectations, the development process should make the next question explicit:

What new information justifies changing the original interpretation?

Earlier evidence only creates value if there is a structured and traceable way to determine how that evidence influences the next development gate.

GSK / depemokimab - how much is the next experiment worth?

Depemokimab presents almost the opposite problem.

Rather than determining whether a program with weakening evidence should stop, GSK needed to determine whether a promising program still required a conventional Phase II dose-ranging step.

The program benefited from an unusually mature evidence base. IL-5 biology was well established; GSK had substantial prior clinical experience with mepolizumab; blood eosinophils provided an informative pharmacodynamic measure; and Phase I data could be interpreted through an established PK/PD relationship.

GSK integrated this information using MIDD and a quantitative decision-making framework. Advancement criteria were defined prospectively using Minimum and Target Values informed by prior mepolizumab data. A Bayesian PK/PD model was then used to estimate whether the proposed depemokimab regimen would achieve the intended pharmacodynamic response. [6]

The development team then addressed a different question: how much information would a conventional Phase IIb efficacy dose-ranging study actually add to the Phase III dose decision?

According to the published simulations, such a study had less than a 3% probability of providing a more precise estimate for the Phase III regimen than the pharmacology-based approach. The authors estimated that the strategy reduced the development timeline by approximately two to three years. [6]

The generalizable lesson is not that Phase II can routinely be bypassed. Depemokimab had unusually favorable conditions: established biology, previous class experience, an informative biomarker, an existing safety evidence base and substantial quantitative-development capability.

The value of the next experiment should be judged by how much it is expected to improve the next decision.

That is fundamentally different from asking whether another study can be performed. It is a Value-of-Information question.

Roche/Genentech / IMbrave151 - extracting information before the endpoint matures

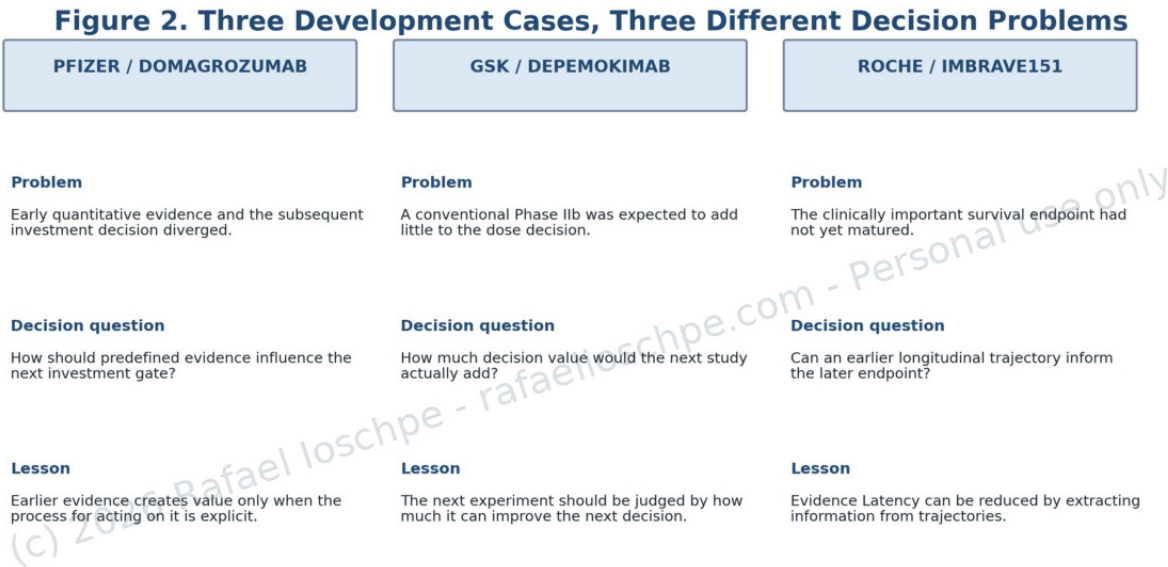
The randomized Phase II IMbrave151 trial in advanced biliary tract cancer presented a third problem: overall survival had not yet matured sufficiently to support the desired development decision.

Investigators applied a previously developed tumor-growth-inhibition-overall-survival model to early longitudinal tumor dynamics. At an interim analysis with 98 patients and approximately 27 weeks of follow-up, the model estimated an OS hazard ratio of 0.74. When survival data later matured, the observed hazard ratio was 0.76. The authors described this as the first prospective use of their TGI-OS framework to support a Phase III development gate. [7]

The innovation was not another early look at the same endpoint. It was the use of an earlier longitudinal signal to update expectations for an outcome that had not yet matured.

When the definitive endpoint is late, which earlier observations can credibly update the probability that the program will ultimately succeed?

This is where reducing Evidence Latency becomes an operational strategy rather than simply a new expression.



The common objective is earlier discrimination among programs that should STOP, PIVOT, LEARN, ENRICH, CONTINUE or ACCELERATE.

Start with the decision, not with the AI

Before choosing an algorithm, a development program should define what must remain true for continued investment and what evidence could materially weaken those assumptions.

This principle has a clear scientific lineage.

Falsifiability is a well-established concept in the philosophy of science, most prominently associated with Karl Popper: a scientific hypothesis must, in principle, be open to empirical refutation. Pharmaceutical development has long applied related ideas using more familiar terminology - proof-of-mechanism, fast-fail, quick-kill, Go/No-Go criteria, futility analyses and stage gates. The NIMH Fast-Fail program, for example, was explicitly designed around biomarker-based proof-of-mechanism in early clinical development. [8,9]

Rather than describe the entire approach as a "Falsification Framework," a more operational construct is:

Prospective Evidence-to-Decision Framework

Falsifiability is one of its principles. The framework links each critical development assumption prospectively to evidence capable of changing the development response.

Figure 3. Prospective Evidence-to-Decision Framework

Critical questions, evidence requirements and potential decision responses are defined prospectively.

Critical development question	Decision-relevant evidence	Potential decision response
Is the target biologically relevant?	Human genetics, functional biology, translational evidence	REASSESS TARGET / STOP
Can the molecule reach the relevant compartment?	PK, PBPK, tissue exposure	DOSE OR MOLECULE PIVOT
Is there sufficient target engagement?	Receptor occupancy, PD marker, imaging	DOSE PIVOT / STOP
Does target engagement alter downstream biology?	Omics, pathway biomarkers, functional measures	REASSESS MECHANISM / PIVOT
Is the early patient trajectory changing as expected?	Imaging, ctDNA, laboratory or digital trajectories	LEARN / ENRICH / STOP
Does the emerging trajectory support later clinical benefit?	Disease models, Bayesian predictive probability	CONTINUE / LEARN / STOP
Will the expected profile remain competitive?	TPP, external evidence, evolving standard of care	REPOSITION / STOP
Would another study materially improve the decision?	Value-of-Information analysis and simulation	LEARN / ACCELERATE

The value of the framework lies in defining the critical questions, evidence requirements and potential decision responses before the answer is known.

Only then should the organization ask where AI can improve the process.

Four practical ways to reduce Evidence Latency

Simply collecting more biomarkers does not constitute a new development strategy. A more useful approach is to design the program around the earliest information capable of materially changing confidence in eventual success.

This process can be organized around four complementary strategies.

1. Endpoint Back-Translation

Development often begins with the ultimate clinical outcome and waits until that outcome becomes measurable.

Endpoint Back-Translation reverses that logic.

Start with the outcome required for the program to succeed and work backward through the causal chain. In the sequence below, the arrows indicate the direction of back-translation, not the biological causal direction:

Required late clinical outcome → required earlier clinical trajectory → required biological response → required pathway modulation → required target engagement → required exposure

The key question becomes:

Which is the earliest link whose failure would materially reduce confidence in the final outcome?

Depemokimab illustrates part of this logic. The dosing decision did not require every downstream clinical question to be resolved because prior evidence allowed the exposure-pharmacodynamic relationship to carry substantial information for that specific Context of Use.

Endpoint Back-Translation does not assume that an earlier signal can substitute for the definitive clinical endpoint. Its purpose is to identify the earliest measurable requirement in the causal chain whose failure would materially weaken the later clinical hypothesis.

2. Map the Falsifiability Horizon

Each critical development assumption has a different earliest point at which it can be tested meaningfully.

This article refers to that interval as the Falsifiability Horizon:

the earliest point at which a critical program assumption can be evaluated with evidence sufficiently informative to influence a development decision.

The horizon varies across the causal chain: exposure may become informative within days, target engagement within days or weeks, biological trajectories over weeks or months, while definitive clinical efficacy may require months or years.

Importantly, Falsifiability Horizon is assumption-specific, not only disease-specific. A therapeutic area may have a long horizon for definitive clinical efficacy and a much shorter one for exposure, target engagement, pathway modulation or another upstream question.

Trontinemab illustrates this distinction in Alzheimer's disease, where meaningful cognitive benefit remains a late clinical readout. In its Phase Ib/IIa Brainshuttle AD study, Roche observed rapid, dose-dependent amyloid-PET reduction together with changes in downstream CSF and plasma biomarkers, including pTau217, and announced progression to Phase III based on the totality of data. Roche later described accelerated Phase III decision-making based on biomarker proof-of-concept as part of a fast-track strategy estimated to shorten time to filing by 21 months; modeling from the Phase Ib/IIa study also informed the Phase III dosing regimen. The clinical Falsifiability Horizon therefore remained long, while upstream exposure, amyloid-removal and biomarker-response questions became decision-relevant earlier. Reducing Evidence Latency does not require making the definitive endpoint mature sooner; it can mean moving the decision upstream in the causal chain when the Context of Use allows it. [16-18]

Development planning should therefore consider not only conventional study milestones, but also the earliest point at which each critical assumption becomes meaningfully testable.

3. Optimize the next readout for information value

The next analysis should not occur only because the protocol specifies "Week 12."

Which next observation is most likely to move the decision?

Depending on the program, that could be receptor occupancy, a PK/PD relationship, ctDNA, imaging, a molecular signature, a digital measurement - or simply additional follow-up.

An observation may be scientifically interesting while having little decision value. Conversely, a relatively simple measurement can be extremely valuable if it materially changes the probability of crossing a predefined development boundary.

The GSK case makes this distinction concrete. The issue was not whether a Phase II study would generate additional data; it clearly would. The question was whether those data were sufficiently likely to improve the decision to justify obtaining them.

4. Look for Multi-Signal Convergence

The future of early development is unlikely to depend on finding one universal "magic biomarker."

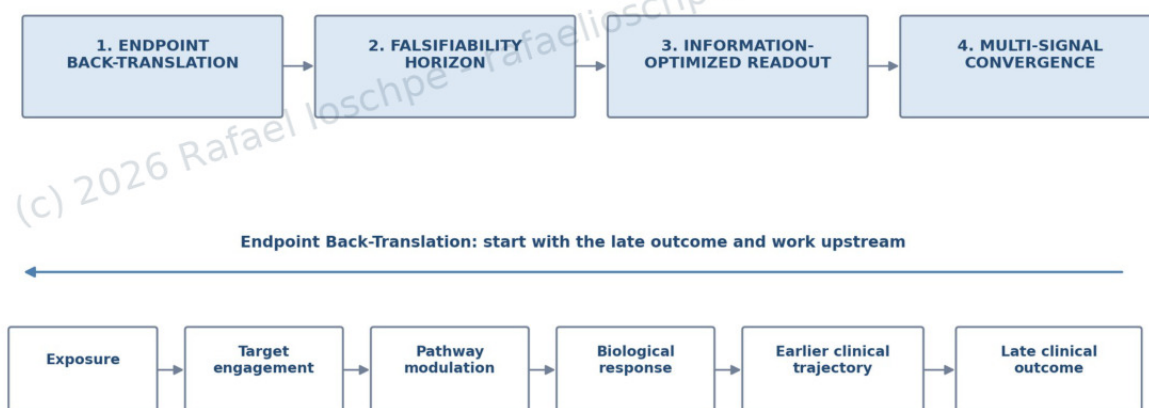
A more realistic opportunity is to determine whether partially independent signals are converging toward - or away from - the development hypothesis. Relevant signals could include:

- exposure
- target engagement
- pathway modulation
- imaging
- molecular trajectories
- early clinical measurements

No single signal may be sufficient. Collectively, however, they may materially alter confidence in the causal chain.

This is one area where AI may provide genuinely new capability: integrating heterogeneous, longitudinal and partially independent signals at a scale that is difficult to reproduce through serial human review.

Figure 4. Four Practical Levers to Reduce Evidence Latency



Note: The lower sequence is shown in the biological causal direction; the upper arrow indicates the direction of Endpoint Back-Translation.

Evidence Latency is reduced when development is designed around the earliest informative decision signal - not simply around more frequent analyses.

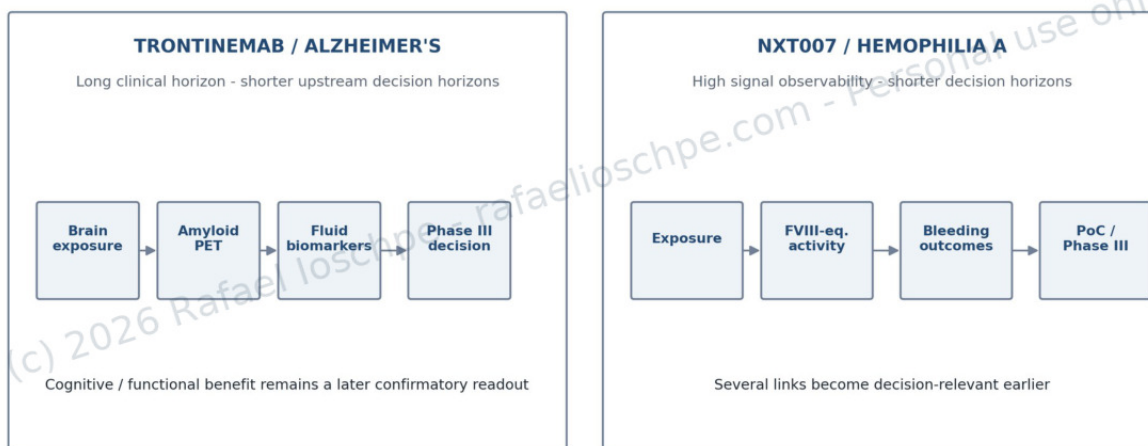
Applicability depends on the biology of the therapeutic area

The decision architecture can be applied broadly, but its ability to generate early decision-grade evidence depends on how observable the causal chain is and how long critical assumptions take to become testable. Signal Observability describes how directly and frequently the sequence between drug exposure and clinical benefit can be measured, while Falsifiability Horizon describes the earliest point at which a specific assumption becomes decision-relevant. Because these properties differ across mechanisms and endpoints, applicability varies substantially across and within therapeutic areas.

Figure 5. Relative Applicability by Therapeutic Area

Therapeutic area	Relative early-decision applicability	What can often be observed earlier	Principal constraint
Oncology	High-very high	ctDNA, imaging, tumor kinetics, molecular signals	Early signals are not automatically validated surrogates
Anti-infectives	High	PK/PD, pathogen susceptibility/load, microbiological response	Pathogen and resistance heterogeneity
Immunology / dermatology	Moderate-high	PD biomarkers, cytokines, tissue, short trajectories	Mechanism- and indication-specific predictiveness
Metabolic disease	Moderate-high	glucose, HbA1c, lipids, exposure-response	Intermediate markers may not capture all long-term outcomes
Rare diseases	Variable / potentially high value	natural history, progression models, selected biomarkers	Sparse data and transportability
Cardiovascular	Moderate	mechanistic biomarkers, imaging, physiological measures	Hard outcomes may remain late
CNS / neurodegeneration	Low-moderate (clinical); higher upstream	target engagement, PET, fluid biomarkers, pathway measures	Long clinical horizons; uncertain biomarker-clinical translation
Hemophilia / selected coagulation disorders	High-very high (selected mechanisms)	exposure, FVIII-equivalent activity, bleeding outcomes	Depends on strength and validity of mechanistic/PD linkage
Psychiatry	Low-moderate	proof-of-mechanism, selected functional measures	Mechanistic evidence may precede reliable efficacy prediction

Figure 6. Same Framework, Different Observable Causal Chains



The contrast shown in Figure 6 illustrates why the horizon should not be assigned at therapeutic-area level alone. In hemophilia A, the Phase I/II NXTAGE program for NXT007 linked dose-dependent exposure to predicted FVIII-equivalent activity and bleeding outcomes. Predicted FVIII-equivalent activity reached the nonhemophilic range from cohort B2 onward, while the mean annualized treated bleeding rate was 0.00 in

the B3 and B4 cohorts. Chugai subsequently advanced the program into Phase III, with two Phase III studies initiated by July 2026. Compared with Alzheimer's disease, several links in this causal chain are more directly observable, allowing multiple development assumptions to become decision-relevant earlier. [19,20]

The NIMH Fast-Fail experience illustrates the distinction. In psychiatry, where reliable early efficacy prediction can be difficult, the development question was shifted toward proof-of-mechanism and whether target engagement altered the hypothesized neural circuitry. [9]

When efficacy cannot be tested early, mechanism sometimes can.

AI does not remove these biological limitations. It cannot create an informative early signal when the biology does not make one observable.

Where AI actually adds value

The phrase "use AI" is too broad to guide a development strategy. Different decision problems require different AI architectures.

Evidence AI

LLMs combined with retrieval-augmented generation and governed knowledge bases can search and connect internal study reports, external literature, competitor results, regulatory evidence and biological knowledge.

The relevant question is not merely what supports a program, but what new evidence has emerged that materially weakens one of its critical assumptions.

This creates the possibility of a continuously operating Scientific Red Team - not empowered to stop a program, but designed to reduce the risk that contradictory evidence is systematically overlooked.

Pattern AI

Longitudinal and multimodal machine-learning methods can examine trajectories across imaging, laboratory data, ctDNA, omics, digital measurements and clinical observations.

These systems are particularly relevant to Multi-Signal Convergence. Their purpose is not simply to maximize predictive accuracy, but to determine whether independent signals are coherently changing the interpretation of the development hypothesis.

Predictive AI

AI/ML models are already being developed to predict Phase II outcomes, clinical-trial termination and eventual drug approval. Aliper et al., for example, described a multimodal platform incorporating target biology, trial design and other information to predict clinical-trial outcomes. [10]

These models can provide useful portfolio intelligence. However, predicting failure is not equivalent to understanding why confidence in a program is declining.

A study may terminate for operational, financial, strategic, safety or efficacy reasons. A prediction model may also learn relationships from sponsor characteristics or study design that are genuinely predictive without directly interrogating the therapeutic mechanism.

For that reason, a single opaque AI-derived Probability of Success should not become the sole decision output.

The underlying confidence structure matters.

Figure 7. Decision Confidence Should Be Decomposed



A composite Probability of Success may remain useful as a summary. It should not obscure which component changed, by how much, and why.

Hybrid mechanistic + AI models

For high-stakes clinical-development decisions, hybrid architectures may be particularly attractive.

PK/PD, PBPK, disease-progression and quantitative-systems-pharmacology models encode biological and pharmacological structure. AI can add pattern recognition, nonlinear relationships and additional dimensions without necessarily discarding mechanistic interpretability.

This direction is consistent with the broader AI-MIDD literature, which increasingly argues for complementarity between data-driven and model-informed methods rather than replacement of one by the other. [1-4]

For a high-impact development decision, interpretability can be an asset rather than a technological limitation.

Decision Latency requires an operating model

Once sufficiently informative evidence exists, the nature of the problem changes.

Because Decision Latency is primarily a process and governance problem, reducing it requires explicit operating mechanisms for translating evidence into action. Four are particularly relevant: predefined decision boundaries, structured overrides, independent challenge and mandatory re-review.

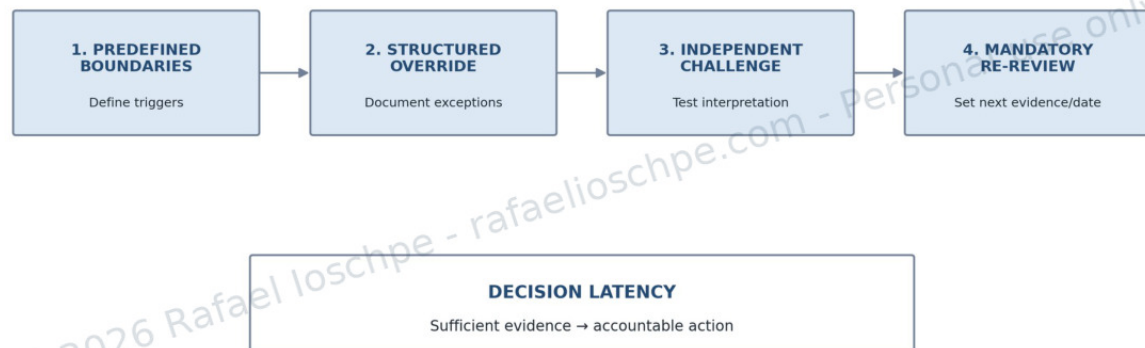
Predefined decision boundaries establish prospectively which forms of evidence should trigger reconsideration of the program and which potential responses - STOP, PIVOT, LEARN, ENRICH, CONTINUE or ACCELERATE - should formally enter the discussion.

Structured override preserves scientific judgment without allowing quantitative criteria to disappear once the result becomes uncomfortable. Leadership may legitimately depart from a prospectively defined rule, but the evidence and rationale supporting that departure should become explicit.

Independent challenge introduces a function capable of testing the interpretation without being directly dependent on the program's progression. Its role is not to advocate for termination, but to challenge assumptions, alternative explanations and selective interpretations of the evidence.

Mandatory re-review prevents LEARN or CONTINUE from becoming indefinite states. When meaningful uncertainty remains, the organization should define which new information is expected, when it should become available and when the next formal decision will occur.

Figure 8. Four Operating Mechanisms to Reduce Decision Latency



AI may accelerate evidence availability, but decision rights, exception handling and review discipline determine how quickly an organization acts. Together, these mechanisms address Decision Latency as what it primarily is: a process and governance problem, ensuring that unresolved uncertainty returns to a defined decision point.

Traceability may become an advantage, not only a constraint

The need for traceability is not unique to drug-development AI, and experience elsewhere in healthcare provides a useful caution.

In a widely deployed U.S. population-health algorithm, Obermeyer and colleagues found that healthcare cost had been used as a proxy for health need. Because spending reflected differences in access and utilization, Black patients with the same algorithmic risk score were, on average, substantially sicker than White patients. The authors estimated that correcting the bias would increase the proportion of Black patients identified for additional care from 17.7% to 46.5%. [11]

This was not a drug-development case, and it should not be interpreted as one. Its relevance is methodological: a model can be technically predictive while still optimizing a proxy that is misaligned with the real decision objective.

For clinical development, that means traceability is not merely a compliance requirement. It is part of determining whether a model is answering the right question, using the right evidence and supporting the intended Context of Use.

Current regulatory thinking is moving in the same direction. FDA's draft AI guidance uses a risk-based credibility framework tied to a clearly defined Context of Use, while the joint FDA-EMA Good AI Practice principles emphasize human-centric design, data governance and documentation, performance assessment and lifecycle management. Final ICH M15 similarly establishes a harmonized framework for

planning, evaluating and documenting MIDD evidence, including AI/ML approaches within its scope. [12-14]

These principles argue against:

data → black-box AI → STOP

A more defensible architecture is:

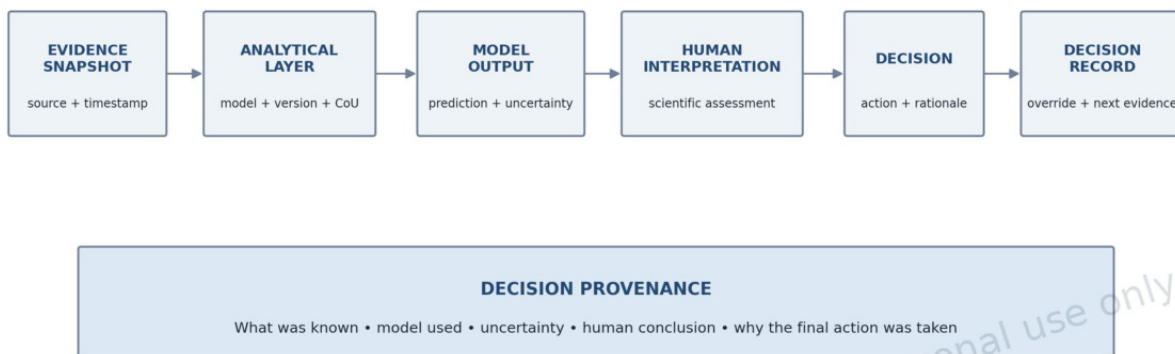
data → traceable evidence → fit-for-purpose model → prediction + uncertainty → scientific interpretation → accountable decision

Properly designed AI may therefore make some development decisions more “reconstructable”, not less.

A conventional portfolio decision may emerge from multiple analyses, presentations, expert discussions and management judgment. Years later, reconstructing exactly what was known and why a particular decision was made can be difficult.

An AI-augmented process can preserve Decision Provenance.

Figure 9. Decision Provenance



The objective is not to eliminate judgment. It is to make judgment reconstructable.

Implementation should start with decision maturity - not AI maturity

A company does not need a large AI organization to begin improving these decisions.

The logical sequence is more likely to be:

Decision discipline → Quantitative development → Integrated evidence → AI augmentation → Portfolio intelligence

The importance of organizational maturity is increasingly supported by industry evidence. A 2026 Pistoia Alliance analysis drawing on workshops, executive surveys and AI/ML use cases concluded that organizational capability and project maturity were important determinants of successful AI deployment, while reproducibility, explainability and governance remained relevant constraints in higher-risk applications. [15]

This suggests that the relevant capability is broader than AI maturity.

It is Quantitative Development Maturity.

That capability sits across Clinical Development, Translational Medicine, Clinical Pharmacology, Pharmacometrics, Biostatistics, Data Science, Regulatory, Quality and Portfolio Governance.

An organization considering implementation could begin with a limited number of programs in which the biological chain is reasonably observable, historical evidence exists, critical decisions can be explicitly defined and model outputs can initially operate in shadow mode, without influencing active decisions.

Only after demonstrating calibration, stability and genuine decision utility should AI progress into assisted decision-making.

Earlier discrimination enables better portfolio prioritization

Earlier discrimination creates strategic value by helping an organization recognize sooner when evidence supporting a program is weakening, when uncertainty can be resolved through targeted learning, and when converging evidence justifies greater commitment. The purpose is not simply to reach a faster stop/go decision; it is to improve allocation of capital, development capacity and management attention across the portfolio.

Weakening evidence → REASSESS / PIVOT / STOP; reallocate resources
Material uncertainty → LEARN / ENRICH before further commitment
Converging evidence → PRIORITIZE / CONTINUE / ACCELERATE

Together, the Pfizer, GSK and Roche cases illustrate three distinct development challenges: how emerging evidence should influence the next investment gate, whether an additional study is worth conducting, and whether an earlier trajectory can meaningfully inform a later endpoint. AI can potentially strengthen each of these processes, but only after the underlying development questions and decision criteria have been made explicit.

From an evidence industry to a decision-ready organization

AI is increasing Pharma's capacity to generate, connect and interrogate evidence. The framework explored here suggests, however, that evidence generation alone is unlikely to become the decisive competitive advantage.

A decision-ready organization would first identify the assumptions on which a development program depends, determine how early each assumption can become meaningfully testable, define which evidence could alter the development path, and establish how much additional information is worth acquiring before the next decision.

Quantitative methods and AI can then be deployed where they create measurable value: shortening the path from an emerging signal to decision-grade evidence, integrating information that would otherwise remain fragmented, updating predictions as new evidence accumulates and continuously challenging the assumptions on which the program depends.

Governance completes that architecture. The final decision remains human, while the evidence, uncertainty, analytical contribution and rationale remain reconstructable through Decision Provenance.

This brings the argument back to a conclusion I reached in an earlier article, The Pharma Market doesn't have an AI problem. It probably has a process problem.

The same issue appears here in a more specific form.

If clinical-development workflows, decision rights, evidence flows, review processes and governance remain unchanged, AI may accelerate individual analyses without materially improving the quality or timing of the development decision. Implementing AI in clinical development is therefore not only a technology challenge. It is also a change-management challenge: processes, incentives, responsibilities and decision policies must evolve with the tools.

This does not diminish the potential of AI. It clarifies where that potential can actually be realized.

The opportunity is not to let AI decide earlier. It is to build an organization capable of reaching the right decision earlier, with stronger evidence, explicit accountability and a clear understanding of how and why that decision was reached.

In that model, AI becomes what it should be: an accelerator of evidence and decision quality - not a substitute for scientific judgment.

And if Pharma is indeed becoming an evidence industry, the competitive advantage may ultimately belong not to the company producing the most evidence, but to the one capable of determining which evidence matters, recognizing it earlier, acting on it at the right moment - and doing so knowing why.

References

1. Barrett JS, Goyal RK, Gobburu J, Baran S, Varshney J. An AI Approach to Generating MIDD Assets Across the Drug Development Continuum. *AAPS J.* 2023;25:70. doi:10.1208/s12248-023-00838-x. <https://pubmed.ncbi.nlm.nih.gov/37430126/>
2. Terranova N, Renard D, Shahin MH, et al. Artificial Intelligence for Quantitative Modeling in Drug Discovery and Development: An Innovation and Quality Consortium Perspective on Use Cases and Best Practices. *Clin Pharmacol Ther.* 2024;115:658-672. doi:10.1002/cpt.3053. <https://pubmed.ncbi.nlm.nih.gov/37716910/>
3. Raman K, Kumar R, Musante CJ, Madhavan S. Integrating Model-Informed Drug Development With AI: A Synergistic Approach to Accelerating Pharmaceutical Innovation. *Clin Transl Sci.* 2025;18:e70124. doi:10.1111/cts.70124. <https://pubmed.ncbi.nlm.nih.gov/39797502/>
4. Mao B, Gao Y, Xu C, et al. Evaluating the Impact of AI-Based Model-Informed Drug Development: A Comparative Review. *AAPS J.* 2025;27:102. doi:10.1208/s12248-025-01075-0. <https://link.springer.com/article/10.1208/s12248-025-01075-0>
5. Fernando K, et al. Achieving end-to-end success in the clinic: Pfizer's learnings on R&D productivity. *Drug Discov Today.* 2022;27:697-704. <https://pmc.ncbi.nlm.nih.gov/articles/PMC8719639/>
6. Zecchin C, Schalkwijk S, Pouliquen IJ, et al. Straight to Phase III: Model-Informed Approach Speeds Depemokimab Clinical Development in Interleukin-5-Driven Diseases. *Clin Pharmacol Ther.* 2026;119:1005-1015. doi:10.1002/cpt.70183. <https://pubmed.ncbi.nlm.nih.gov/41556088/>
7. Shemesh CS, Chan P, Marchand M, et al. Early Decision Making in a Randomized Phase II Trial of Atezolizumab in Biliary Tract Cancer Using a Tumor Growth Inhibition-Survival Modeling Framework. *Clin Pharmacol Ther.* 2023;114:644-651. doi:10.1002/cpt.2953. <https://pubmed.ncbi.nlm.nih.gov/37212707/>
8. Popper KR. *The Logic of Scientific Discovery.* London: Hutchinson; 1959.
9. Krystal AD, Pizzagalli DA, Smoski M, et al. A randomized proof-of-mechanism trial applying the fast-fail approach to evaluating kappa-opioid antagonism as a treatment for anhedonia. *Nat Med.* 2020;26:760-768. doi:10.1038/s41591-020-0806-7. <https://pubmed.ncbi.nlm.nih.gov/32231295/>
10. Aliper A, Kudrin R, Polykovskiy D, et al. Prediction of Clinical Trials Outcomes Based on Target Choice and Clinical Trial Design with Multi-Modal Artificial Intelligence. *Clin Pharmacol Ther.* 2023;114:972-980. doi:10.1002/cpt.3008. <https://pubmed.ncbi.nlm.nih.gov/37483175/>
11. Obermeyer Z, Powers B, Vogeli C, Mullainathan S. Dissecting racial bias in an algorithm used to manage the health of populations. *Science.* 2019;366:447-453. doi:10.1126/science.aax2342. <https://pubmed.ncbi.nlm.nih.gov/31649194/>

12. U.S. Food and Drug Administration. Considerations for the Use of Artificial Intelligence to Support Regulatory Decision-Making for Drug and Biological Products. Draft Guidance. January 2025.
<https://www.fda.gov/regulatory-information/search-fda-guidance-documents/considerations-use-artificial-intelligence-support-regulatory-decision-making-drug-and-biological>
13. FDA / EMA. Guiding Principles of Good AI Practice in Drug Development. January 2026.
<https://www.fda.gov/about-fda/artificial-intelligence-drug-development/guiding-principles-good-ai-practice-drug-development>
14. International Council for Harmonisation / U.S. Food and Drug Administration. M15 General Principles for Model-Informed Drug Development. Final Guidance. June 2026. <https://www.fda.gov/regulatory-information/search-fda-guidance-documents/m15-general-principles-model-informed-drug-development>
15. Makarov VA, Sanger M, Saiz R, Hartung M, Lynch N. Why Do AI Projects Fail in Drug Development and Pharma? Insights from Multi-Year Pistoia Alliance Studies. J Comput Aided Mol Des. 2026.
<https://link.springer.com/article/10.1007/s10822-026-00894-3>
16. Roche. Roche presents novel therapeutic and diagnostic advancements in Alzheimer's at AD/PD 2025. 3 April 2025. <https://www.roche.com/investors/updates/inv-update-2025-04-03>
17. Roche. Pharma Day 2025 - Embrace ambitious R&D objectives. Trontinemab fast-track program: accelerated Phase III decision-making based on biomarker proof-of-concept; 21 months faster to filing. September 2025. https://assets.roche.com/f/176343/x/059c686d27/20250922_pharma-day-2025_vf_online.pdf
18. Roche. Roche presents new data in Alzheimer's disease from across its integrated pharmaceutical and diagnostics portfolio at AAIC. 7 July 2026. Modeling from the Phase Ib/Ila Brainshuttle AD study informed dosing for Phase III TRONTIER 1 and 2. <https://www.roche.com/investors/updates/inv-update-2026-07-07>
19. Nogami K, You C-W, Park Y-S, et al. NXTAGE: a phase 1/2 study of NXT007 to assess safety, pharmacokinetics, and efficacy in hemophilia A without inhibitors. Blood. 2026;147(19):2261-2271. doi:10.1182/blood.2025031894. <https://pubmed.ncbi.nlm.nih.gov/41643195/>
20. Chugai Pharmaceutical Co., Ltd. Chugai Announces 2026 2nd Quarter Results. 24 July 2026. Two Phase III studies for NXT007 initiated.
https://www.chugai-pharm.co.jp/english/news/detail/20260724160000_1265.html