Available online www.ijpras.com

International Journal of Pharmaceutical Research & Allied Sciences, 2025, 14(4):78-88
https://doi.org/10.51847/XX6cXGgYPD

Original Article

ISSN : 2277-3657
CODEN(USA) : IJPRPM

Why Promising Pharmaceutical Technologies Stall between Laboratory Validation and Clinical Translation: A Failure-Mechanism Framework for Pharmaceutical Innovation

Yifan Zhao[1]*, Liang Chen[1], Ahmad Fauzi[2], Nor Hayati[3]

[1]*Department of Pharmaceutical Translation and Failure Mechanisms, Faculty of Pharmacy, China Agricultural University, Beijing, China.*
[2]*Department of Laboratory-to-Clinic Translation, Faculty of Pharmacy, University of Malaya, Kuala Lumpur, Malaysia.*
[3]*Department of Innovation Barriers and Implementation, Faculty of Pharmacy, Universiti Putra Malaysia, Serdang, Malaysia.*

***Email:** yifan.zhao@cau.edu.cn*

ABSTRACT

Promising pharmaceutical technologies frequently demonstrate reproducible technical performance under laboratory conditions yet fail to progress toward patient-relevant clinical use. Such attrition is often attributed to isolated scientific, regulatory, financial, or operational obstacles, obscuring the relationships through which weaknesses in one domain propagate across development. This Original Failure-Mechanism Framework Article proposes a multidimensional architecture for explaining translational stalls beyond technical performance. The framework distinguishes laboratory validation from translation readiness and organizes failure mechanisms into six interacting domains: reproducibility and model-system failure; mechanistic and biological-validity failure; manufacturing and quality-system failure; safety and pharmacological translation failure; clinical-workflow and implementation failure; and financing, coordination, and incentive misalignment. It further introduces three proposed constructs: the failure-mechanism profile, which represents the configuration of barriers affecting a technology; evidence debt, which arises when decision-critical assumptions remain unresolved across development transitions; and the translation observatory, a cross-functional mechanism for tracing evidence, uncertainty, and corrective action. Diagnostic indicators and corrective strategies are organized around claim specification, context of use, assumption exposure, falsification testing, and reassessment of decision readiness. The framework does not assign universal readiness scores or imply that satisfying one domain compensates for failure in another. Its relationships require retrospective evaluation in stalled programmes, prospective comparison across technology classes, inter-rater assessment, and testing of whether mechanism-based diagnosis improves decisions beyond existing development practices. The contribution is therefore a proposed conceptual synthesis for structured inquiry, not a validated predictive model, clinical recommendation, regulatory determination, or deployment-ready system.

Key words: *Translational pharmaceutical research, Translation readiness, Reproducibility, Manufacturing feasibility, Clinical relevance, Evidence maturity*

Received: 12 August 2025; **Revised:** 20 November 2025; **Accepted:** 24 November 2025

INTRODUCTION

Pharmaceutical innovation repeatedly produces technologies that appear promising in laboratory assays but do not become clinically useful products. Clinical-stage failure can arise from interacting deficiencies in efficacy,

safety, exposure, target selection, study design, and patient stratification rather than from one readily identifiable technical defect [1]. Laboratory performance therefore represents evidence about a bounded experimental claim, not proof that the underlying technology can survive the scientific and organizational transitions required for clinical translation.

Attrition is also heterogeneous. Estimated development success varies across clinical phases, therapeutic areas, indications, and analytical assumptions, making a single generalized failure rate an inadequate description of translation [2]. A technology can fail before human testing, after apparently favorable early clinical findings, during process transfer, or when clinical and organizational requirements reveal that its intended use is impractical. These pathways may share an outcome while reflecting fundamentally different mechanisms.

Multidomain development frameworks have emphasized that target biology, tissue exposure, safety, patient selection, and value must be considered jointly [3]. Nevertheless, many development systems continue to reward successful completion of local tasks: an assay produces a signal, a model shows activity, a formulation meets an initial specification, or a prototype functions under controlled conditions. Such achievements may be scientifically meaningful while remaining insufficient for the next translational decision.

This article develops an original, non-empirical Failure-Mechanism Framework for explaining why local validation may not mature into translation readiness. It defines six interacting stall-mechanism domains and proposes diagnostic constructs, evidence relationships, and decision boundaries for investigating them. The framework is intended to organize questions and testable propositions; it is not presented as empirically validated, universally applicable, regulator-endorsed, or capable of predicting the fate of an individual programme.

Translational attrition beyond technical performance

Technical performance concerns whether a technology produces a prespecified output under defined conditions. Translation readiness concerns whether the available evidence is sufficiently reproducible, biologically relevant, manufacturable, pharmacologically credible, usable, and organizationally supported for a specified next decision. Analyses of therapeutic mechanisms and indications show that extensive discovery activity can coexist with poorly differentiated or weakly productive mechanism–indication choices [4]. Technical output and translational value are therefore related but non-equivalent constructs.

This distinction matters because increasing experimental speed or project throughput cannot compensate for poor decision relevance. Pharmaceutical productivity depends partly on which hypotheses enter development, how uncertainty is resolved, and whether programmes are terminated or redirected when critical assumptions fail [5]. A technology may consequently stall despite efficient experimentation when the experiments optimize a surrogate, model, material, or workflow that is disconnected from the intended pharmaceutical consequence.

Attrition also has economic consequences, but these consequences should not be reduced to one universal development-cost estimate. Estimates of investment required to bring medicines to market vary with company samples, accounting assumptions, treatment of failed projects, development duration, and opportunity cost [6]. Financial exposure establishes the importance of earlier diagnosis, yet high expenditure does not reveal whether a programme's primary weakness is biological, manufacturing, pharmacological, operational, or organizational.

The framework therefore treats attrition as a loss of continuity between evidence and decision use. A stall is proposed as a delay, abandonment, repeated redesign, or failure to cross a development transition despite apparently favorable evidence in at least one local domain. A stall mechanism is the causally plausible process through which that local evidence becomes insufficient, misleading, non-transferable, or unusable. This interpretation does not assume that all delays represent failure or that every terminated programme could have been rescued.

Proposed classification of stall mechanisms

Predictive validity provides the conceptual foundation for classifying stalls. An experimental tool is valuable for translation only insofar as its outputs are associated with the later outcome relevant to the intended decision [7]. A technically precise assay may therefore have low translational value when it measures the wrong phenomenon, operates outside its validated domain, or supports a claim more expansive than its evidence.

The first proposed class is reproducibility and model-system failure, in which an observation cannot be reproduced, is highly sensitive to experimental context, or originates in a system that poorly represents the intended biological state. Weak prior plausibility and false-positive findings can make apparently successful preclinical selection insufficiently informative about human causal biology [8]. This class includes failures of

repeatability, robustness, external validity, model qualification, and correspondence between the experimental system and the intended patient population.

The second and third classes are mechanistic and biological-validity failure and manufacturing and quality-system failure. The former occurs when the observed phenotype does not establish the proposed causal mechanism or patient-relevant biology. The latter occurs when a technology cannot be produced, characterized, controlled, transferred, or maintained in a form representative of its intended product. Emerging human-relevant systems may strengthen parts of the evidence chain, but they remain conditional on qualification, reproducibility, endpoint relevance, and context of use [9].

The remaining classes are safety and pharmacological translation failure, clinical-workflow and implementation failure, and financing, coordination, and incentive misalignment. These concern, respectively, the inability to connect dose, exposure, distribution, toxicity, and therapeutic window; incompatibility with patient needs, professional workflow, infrastructure, or health-system conditions; and organizational arrangements that reward milestone completion while leaving decision-critical uncertainties unresolved. The classes are proposed as interacting and non-exclusive rather than as mutually exclusive causes.

Three additional constructs organize their relationships. A failure-mechanism profile is the technology-specific configuration of primary barriers and interacting dependencies. Evidence debt is the accumulation of unresolved assumptions carried across a decision transition. A translation observatory is a proposed cross-functional function that traces claims, evidence, uncertainties, handoffs, and corrective actions across development. These constructs do not establish numerical readiness levels or replace product-specific scientific, clinical, manufacturing, or regulatory evaluation.

Reproducibility and model-system failure

Reproducibility is not a single property. It includes the capacity to repeat an analysis, reproduce an observation under comparable conditions, and obtain sufficiently robust findings across investigators, sites, materials, populations, or model variants. Reproducible science consequently depends on study design, transparent methods, bias control, reporting, incentives, and mechanisms for cumulative correction [10]. Repetition of an inadequately specified experiment cannot resolve weaknesses in the claim being tested.

Evidence from replication initiatives illustrates that technically plausible preclinical findings may be difficult to reproduce because essential materials, procedures, exclusions, analyses, or contextual conditions are insufficiently reported. Repeated effects may also be smaller or less consistent than those in originating studies [11]. These outcomes do not prove misconduct or invalidate an entire field, but they show that evidence maturity cannot be inferred from publication or initial statistical significance alone.

Excessive standardization creates a different risk. Restricting biological variation can increase precision within one experiment while producing results that are fragile across strains, sexes, laboratories, environments, or procedural histories [12]. Heterogeneity should not be introduced indiscriminately; rather, the relevant sources of variation should be selected according to the intended inferential domain. A model designed for mechanistic isolation may require different validation from one intended to anticipate heterogeneous patient responses.

Model-system failure extends beyond non-replication. A result may be reproducible yet remain non-translatable because the system lacks the relevant cell states, tissue architecture, immune interactions, disease history, exposure conditions, comorbidities, or clinical heterogeneity. The proposed framework therefore separates observation reproducibility from model–decision correspondence. **Figure 1** presents the original conceptual synthesis for failure-mechanism map for stalled pharmaceutical technologies, showing how the article's principal components, evidence relationships, uncertainties, and decision boundaries are connected.

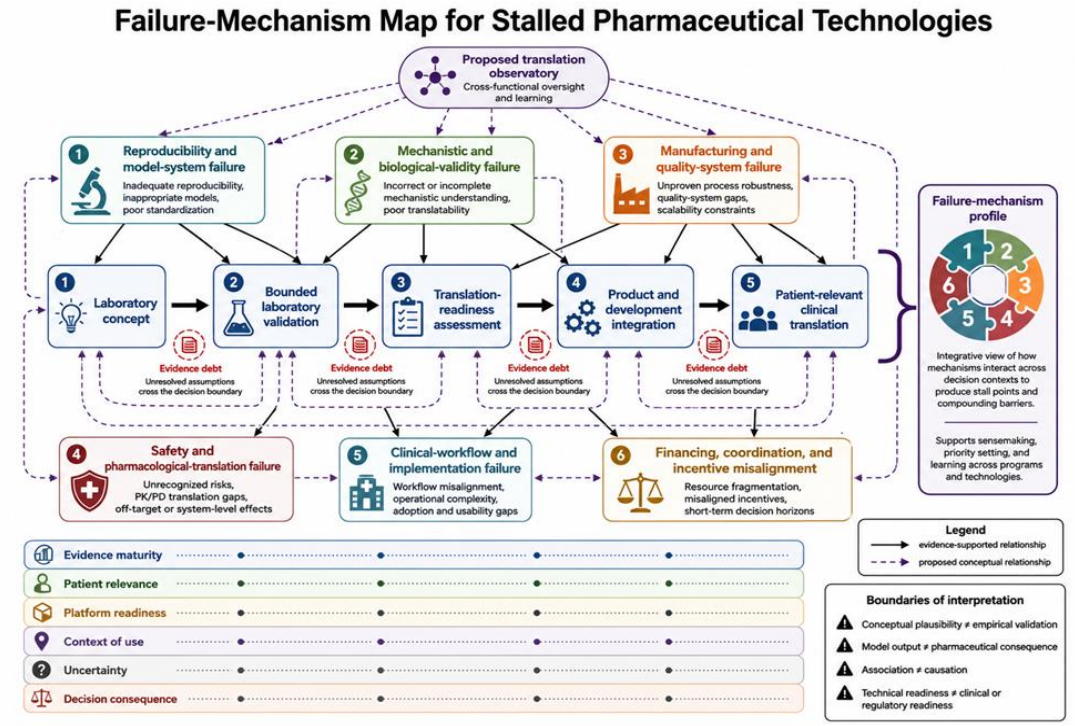

Figure 1. Failure-Mechanism Map for Stalled Pharmaceutical Technologies.

The figure is an original conceptual synthesis developed for "Why Promising Pharmaceutical Technologies Stall between Laboratory Validation and Clinical Translation: A Failure-Mechanism Framework for Pharmaceutical Innovation." Arrows and grouping indicate proposed or evidence-supported relationships rather than measured effect sizes. The figure does not represent a validated predictive model, clinical recommendation, regulatory determination, or deployment-ready system.

The proposed diagnostic sequence for this class begins by specifying the exact claim, its intended decision, and the model features assumed to connect them. Investigators should then test independent replication, relevant heterogeneity, orthogonal endpoints, model perturbation, and transfer across conditions representing the intended domain. Discordance should trigger revision of the claim or model rather than automatic averaging of results. These criteria remain proposed: reproducibility may be necessary for translation, but it cannot establish that a reproducible observation is mechanistically correct, patient-relevant, manufacturable, safe, or clinically useful.

Mechanistic and biological-validity failure

Mechanistic validity concerns whether the intervention acts through the proposed causal pathway rather than merely producing a correlated phenotype. Human genetic support is associated with improved development prospects, but the magnitude of that association depends on target definitions, datasets, and analytical assumptions [13]. Genetic support should therefore strengthen a causal argument without being treated as proof that a particular modality, dose, or intervention will benefit patients.

Human loss-of-function variation can clarify target function and identify biological consequences of reduced activity [14]. However, lifelong genetic exposure differs from pharmacological intervention in timing, magnitude, reversibility, tissue distribution, and compensatory adaptation. Apparent tolerance of genetic loss cannot be translated directly into clinical safety, while disease association does not independently establish that modifying the target will reverse established pathology.

Human organoids can reproduce selected aspects of tissue organization and disease biology, offering an intermediate level of complexity between reductionist assays and intact organisms [15]. Their relevance remains conditional on cellular maturity, microenvironment, vascular and immune representation, experimental reproducibility, exposure conditions, and endpoint selection. Human origin alone does not guarantee patient relevance or superiority for every developmental question.

The framework proposes a mechanistic-concordance test requiring alignment among target engagement, causal perturbation, disease-context relevance, achievable exposure, and reversal of a decision-relevant phenotype. Discordance among these elements should prompt mechanism revision, model replacement, or restriction of the

intended claim. These criteria are investigational and do not constitute validated development gates or evidence of clinical efficacy.

Manufacturing and quality-system failure

A laboratory technology becomes pharmaceutically relevant only when its defining properties can be reproduced in material representative of the intended product. Continuous-manufacturing scholarship illustrates that process integration, equipment behaviour, material flow, analytical control, and scale must be considered together rather than as separable technical tasks [16]. Successful preparation of small laboratory quantities does not establish a controlled or transferable manufacturing process.

Manufacturing readiness also requires traceability, process monitoring, control strategy, material-diversion logic, batch definition, data integrity, and lifecycle verification [17]. These requirements become especially important when process history influences critical material attributes or biological performance. A technology can therefore stall even when its active principle remains scientifically promising because its product state cannot be defined or maintained consistently.

Pharmaceutical quality is not secured through end-product testing alone. It depends on accumulated process knowledge, capable operations, appropriate surveillance, deviation learning, and continual improvement [18]. Specifications can detect some failures but cannot compensate for an inadequately understood process or an analytical method that does not measure the attributes responsible for performance.

The framework consequently proposes a manufacturing–biology coupling rule: biological and pharmacological claims should be tested using material representative of the intended process, control strategy, storage history, and administration configuration. Changes in scale, equipment, raw materials, formulation, or purification should trigger assessment of whether earlier evidence remains applicable. This rule requires product-specific validation and does not establish manufacturability, comparability, or regulatory acceptability.

Table 1 organizes the evidence, constructs, relationships, uncertainties, and boundary conditions required for definitions, components, relationships, and intended uses in why promising pharmaceutical technologies stall between laboratory validation and.

Table 1. Definitions, components, relationships, and intended uses in Why Promising Pharmaceutical Technologies Stall between Laboratory Validation and.

Component or scientific dimension	Problem addressed	Inputs or determinants	Proposed mechanism or relationship	Expected contribution	Evidence required	Failure or uncertainty risk	Boundary statement
Laboratory validation	Local technical claims may be mistaken for translation readiness.	Assay design, endpoint, model, controls, material state	**Proposed:** validation supports only the claim and context directly examined.	Prevents expansion from experimental success to product-level inference.	Reproducible methods, prespecified endpoints, orthogonal confirmation [10]	Precise measurement of a clinically irrelevant phenomenon	Does not establish patient benefit or development readiness.
Predictive validity	Experimental outputs may not correspond to later outcomes.	Reference outcome, context of use, decision consequence	Tool value depends on correspondence with the intended decision [7].	Links assay selection to pharmaceutical consequence.	Prospective or external outcome comparison	Circular validation or inappropriate reference standards	Validity remains bounded to the evaluated use.
Reproducibility and model correspondence	Findings may be unstable or generated in an unrepresentative system.	Biological variation, reporting, laboratory context, model features	**Proposed:** observation reproducibility and model–decision correspondence are separate requirements.	Distinguishes repeatable signals from translatable evidence.	Independent replication and domain-relevant heterogeneity [12]	Reproduction of the same model-specific bias	Reproducibility alone does not establish biological validity.
Mechanistic concordance	Phenotypic change may not demonstrate	Target engagement, perturbation,	**Proposed:** mechanism, exposure, and	Identifies causal gaps before later development.	Human-relevant evidence and causal	Association interpreted as causation	Does not establish efficacy in patients.

	causal therapeutic action.	exposure, disease context	phenotype must align.		perturbation [13]		
Manufacturing readiness	Product properties may drift during scale, transfer, or storage.	Process history, materials, equipment, controls, analytics	**Proposed:** pharmaceutical claims remain applicable only when material remains representative.	Connects process development to biological evidence.	Integrated process and scale evidence [16]	Unmeasured critical attributes or non-representative batches	Does not establish commercial capability or approval.
Quality-system maturity	Testing may detect defects without preventing recurrence.	Process knowledge, surveillance, deviations, lifecycle learning	Quality emerges from capable systems rather than final testing alone [18].	Supports persistent control and corrective learning.	Validated methods, process capability, deviation investigation	Specifications disconnected from performance	Compliance with a test does not prove clinical suitability.
Failure-mechanism profile	Different barriers may produce the same visible stall.	Evidence gaps across all framework domains	**Original proposed construct:** technology-specific configuration of primary and interacting barriers	Directs mechanism-specific investigation.	Retrospective and prospective classification studies	Oversimplification or subjective coding	Not a validated score or prediction algorithm.
Evidence debt and translation observatory	Unresolved assumptions may pass through organizational handoffs.	Claims, decisions, ownership, uncertainty, corrective actions	**Original proposed constructs:** evidence debt records unresolved assumptions; the observatory traces them across transitions.	Improves traceability and cross-functional challenge.	Prospective governance evaluation	Administrative burden or false confidence from documentation	Does not replace scientific, clinical, regulatory, or investment judgment.

Safety and pharmacological translation failure

Safety translation requires more than the absence of toxicity in a selected model. Cross-company evidence indicates that conventional nonclinical programmes can support safe entry into first-in-human studies, while the detectability and clinical correspondence of toxicities vary by organ system and context [19]. Such evidence supports risk management for trial entry but cannot guarantee broader safety, therapeutic benefit, or later-phase success.

Pharmacological translation additionally depends on whether models are credible for their defined context of use. Mechanistic and in-silico models should be evaluated according to verification, validation, applicability, uncertainty, and the consequences of the decision they inform [20]. A model suitable for hypothesis generation may be insufficient for dose selection, safety-margin interpretation, population extrapolation, or product comparison.

The proposed pharmacological-translation test integrates achievable exposure, target engagement, tissue distribution, duration of action, reversibility, therapeutic window, and uncertainty. Each element must be connected to the intended claim and administration context. Confidence in one element cannot compensate automatically for an unresolved deficiency in another, and model credibility must remain bounded to the use for which supporting evidence was generated [20].

Alternative explanations should be considered when translation fails. Lack of clinical activity may reflect inadequate exposure rather than incorrect biology, while toxicity may arise from modality, formulation, metabolite, distribution, or dosing schedule rather than the target itself. The framework organizes these

possibilities but does not select a clinical dose, authorize human exposure, or determine that a programme is safe to proceed.

Clinical-workflow and implementation failure

Clinical utility depends on whether a technology can be integrated into actual care. Health technologies may be rejected, abandoned, or unable to scale because perceived value, adopter burden, organizational capacity, system conditions, and adaptation requirements interact [21]. A technology may therefore remain clinically unused despite analytical validity or biological activity.

Implementation determinants include features of the intervention, inner and outer settings, participating individuals, and implementation processes [22]. For pharmaceutical technologies, these determinants may involve administration complexity, monitoring requirements, diagnostic dependence, professional training, storage, infrastructure, interpretability, reimbursement, access, and accountability. Their importance varies across populations and health systems.

The framework proposes a patient–workflow concordance test requiring alignment between the intended therapeutic benefit and the practical conditions of delivery. Development should examine who performs each action, what infrastructure is required, what burden is imposed, how failures are detected, and whether the intended population can realistically access the technology. These questions should influence product design before clinical evidence becomes difficult or expensive to reinterpret.

Implementation readiness cannot compensate for absent efficacy or unacceptable risk. Conversely, favorable clinical findings do not guarantee adoption, scale, affordability, or sustainability. Clinical-workflow failure is therefore treated as a distinct but interacting mechanism rather than as a late commercial inconvenience. Its diagnostic criteria remain context-dependent and require evaluation with patients, professionals, organizations, and delivery systems.

Table 2 organizes the evidence, constructs, relationships, uncertainties, and boundary conditions required for testable propositions, evidence requirements, and validation criteria for why promising pharmaceutical technologies stall between laboratory validation and.

Table 2. Testable propositions, evidence requirements, and validation criteria for Why Promising Pharmaceutical Technologies Stall between Laboratory Validation and.

Architecture layer or process stage	Core function	Information or material flow	Interaction with other components	Validation criterion	Potential failure mode	Human or experimental responsibility	Readiness boundary
Claim definition	Specify what the technology has demonstrated.	Experimental output → bounded scientific claim	Anchors every later evidence request	Independent reviewers can identify the claim, endpoint, and context.	Vague or expanding claims	Investigators and programme leadership	**Proposed:** no transition when the claim is undefined.
Model qualification	Establish whether the system represents the intended decision domain.	Model characteristics → expected human relevance	Connects reproducibility with biological validity	Replication, perturbation, and domain-relevant comparison [11]	Stable but non-representative model	Experimental scientists and translational specialists	Qualification is specific to context of use.
Mechanistic assessment	Test causal alignment among intervention, target, and phenotype.	Perturbation and engagement evidence → causal interpretation	Informs safety, exposure, and patient selection	Concordant target engagement and decision-relevant phenotype [14]	Association mistaken for therapeutic causation	Biologists, pharmacologists, and disease experts	Does not establish clinical efficacy.
Product realization	Preserve defining attributes through manufacture and storage.	Materials and process history → controlled product	Couples manufacturing with biological performance	Representative material, process control, and transfer evidence [17]	Scale-dependent drift or unmeasured attributes	Process, analytical, formulation, and quality teams	Laboratory preparation is not manufacturing readiness.

Pharmacological translation	Connect dose and exposure to intended benefit and risk.	Dose → exposure → distribution → effect	Links mechanism to safety and trial design	Context-specific model credibility and uncertainty analysis [20]	Exposure mismatch or unsupported extrapolation	Pharmacologists, toxicologists, modellers, clinicians	Does not determine a safe or effective dose.
Clinical implementation	Fit the intervention to patients and care systems.	Product and information → clinical workflow	Interacts with access, infrastructure, monitoring, and value	Contextual assessment across users and settings [22]	Nonadoption, abandonment, or unsustainable burden	Patients, professionals, implementers, organizations	Clinical promise is not implementation readiness.
Evidence-debt control	Expose assumptions before they cross decision gates.	Unresolved assumption → accountable corrective action	Operates across every layer	**Proposed:** prospective traceability and closure of critical assumptions	Milestone passage without uncertainty resolution	Cross-functional decision owners	Documentation alone does not establish evidence maturity.
Translation-observatory evaluation	Detect recurring and interacting stall mechanisms.	Cross-domain evidence → diagnosis → falsification test	Integrates all architecture layers	**Proposed:** reliability, feasibility, incremental decision value, and prospective outcome evaluation	Bureaucratic burden or false certainty	Independent multidisciplinary governance	Not deployment-ready until empirically evaluated.

Financing, coordination, and incentive misalignment

Scientific development occurs within organizations that distribute capital, expertise, authority, and risk. Comparative pharmaceutical analyses indicate that R&D spending and organizational scale relate to aggregate outputs, but expenditure alone does not identify whether a particular programme possesses decision-relevant evidence [23]. Additional funding may accelerate work while preserving the assumption responsible for the stall. Collaboration can increase access to technologies and capabilities, yet value creation depends on aligned objectives, responsibilities, decision points, and expectations [24]. Problems arise when partners use different definitions of success, assign evidence generation to one another, or transfer a programme without transferring the assumptions and limitations attached to its data.

The framework defines evidence debt as unresolved decision-critical uncertainty carried across milestones or organizational handoffs. Misaligned incentives may encourage teams to demonstrate immediate progress rather than falsify foundational claims. Collaboration principles emphasize that responsibility and value expectations must be explicit [24], but evidence debt remains an original proposed construct requiring operational definition and empirical study.

Financing and coordination failures should not become catch-all explanations. A scientifically unsound programme is not rescued by improved governance, and termination does not necessarily indicate organizational dysfunction. The proposed category applies when incentives, ownership, sequencing, or information transfer plausibly prevent appropriate evidence generation, interpretation, correction, or stopping decisions.

Diagnostic indicators and corrective strategies

Diagnosis begins with evidence that can be inspected. Transparent reporting of experimental design, bias controls, sample definition, exclusions, and analysis is necessary for determining what was tested and whether the result can be interpreted [25]. Reporting quality does not itself establish reproducibility or validity, but inadequate reporting can make both impossible to assess.

Corrective evidence should be proportional to the importance of the model, the consequence of the decision, its context of use, and residual uncertainty [26]. High-influence decisions require stronger verification, validation, applicability assessment, and challenge than exploratory choices. The relevant standard is therefore not maximal evidence for every question but adequate credibility for a clearly defined decision.

Proposed diagnostic indicators include effect attenuation across sites, sensitivity to minor procedural changes, discordance among models, failure of mechanistic perturbation, scale-dependent performance drift, exposure mismatch, recurrent reformulation, workflow incompatibility, and repeated deferral of unresolved questions.

Replication evidence demonstrates why attenuation and incomplete procedural traceability deserve attention [11], but the complete indicator set remains an original synthesis.

The corrective sequence is to define the claim, identify the next decision, expose connecting assumptions, classify the likely stall mechanism, select a test capable of falsifying the explanation, and reassess the claim's scope. Risk-informed credibility principles support linking evidence requirements to decision consequence [26]. No indicator creates an automatic stop/go decision, and corrective action may appropriately result in restriction, redesign, postponement, or termination.

Boundary conditions and research priorities

Autonomous experimentation can accelerate iterative search, optimization, and data generation. Self-driving laboratories nevertheless inherit the limitations of their objectives, assays, measurements, search spaces, and decision rules [27]. Rapid optimization of an invalid surrogate may deepen rather than resolve a translational stall. Automation should therefore be evaluated as an experimental capability, not as independent evidence of translation readiness.

Translational science addresses recurring barriers across discovery, development, dissemination, and implementation and seeks solutions that generalize across projects [28]. The proposed framework contributes to this agenda by linking mechanisms that are usually examined separately. It does not claim that every technology follows one pathway or that every stall can be assigned a single dominant cause.

Empirical evaluation should begin with retrospective coding of programmes that progressed, stalled, changed indication, or were terminated. Studies should assess whether mechanism classes can be applied reliably, whether evidence debt can be operationalized, and whether the framework explains decisions beyond conventional stage labels. Prospective studies should test whether framework-guided challenges alter evidence generation, programme redirection, or uncertainty resolution without delaying beneficial development.

Further research should compare modalities, disease areas, development stages, organizational settings, and health systems. Translation-observatory pilots should evaluate feasibility, burden, independence, and unintended effects. Until such work is completed, the framework is appropriate for structured inquiry and hypothesis generation, not success prediction, portfolio ranking, regulatory submission, clinical decision-making, or autonomous programme governance.

CONCLUSION

Promising pharmaceutical technologies can stall because local technical evidence fails to remain reproducible, biologically meaningful, manufacturable, pharmacologically credible, clinically usable, or organizationally actionable as decision contexts change. The proposed Failure-Mechanism Framework integrates these domains while distinguishing laboratory validation from translation readiness. Its original constructs—failure-mechanism profile, evidence debt, and translation observatory—offer a vocabulary for tracing discordance and selecting corrective tests. Their value remains unestablished until the framework is evaluated across real development programmes. The framework should therefore be understood as a bounded conceptual architecture for investigating translational attrition, not as a validated prediction system, clinical recommendation, regulatory standard, universal readiness model, or deployment-ready governance mechanism.

ACKNOWLEDGMENTS: None

CONFLICT OF INTEREST: None

FINANCIAL SUPPORT: None

ETHICS STATEMENT: None

REFERENCES

1. Sun D, Gao W, Hu H, Zhou S. Why 90% of clinical drug development fails and how to improve it? Acta Pharm Sin B. 2022;12(7):3049-62. doi:10.1016/j.apsb.2022.02.002
2. Wong CH, Siah KW, Lo AW. Estimation of clinical trial success rates and related parameters. Biostatistics. 2019;20(2):273-86. doi:10.1093/biostatistics/kxx069

3. Morgan P, Brown DG, Lennard S, Anderton MJ, Barrett JC, Eriksson U, et al. Impact of a five-dimensional framework on R&D productivity at AstraZeneca. Nat Rev Drug Discov. 2018;17(3):167-81. doi:10.1038/nrd.2017.244
4. Shih HP, Zhang X, Aronov AM. Drug discovery effectiveness from the standpoint of therapeutic mechanisms and indications. Nat Rev Drug Discov. 2018;17(1):19-33. doi:10.1038/nrd.2017.194
5. Ringel MS, Scannell JW, Baedeker M, Schulze U. Breaking Eroom's Law. Nat Rev Drug Discov. 2020;19(12):833-4. doi:10.1038/d41573-020-00059-3
6. Wouters OJ, McKee M, Luyten J. Estimated research and development investment needed to bring a new medicine to market, 2009-2018. JAMA. 2020;323(9):844-53. doi:10.1001/jama.2020.1166
7. Scannell JW, Bosley J, Hickman JA, Dawson GR, Truebel H, Ferreira GS, et al. Predictive validity in drug discovery: What it is, why it matters and how to improve it. Nat Rev Drug Discov. 2022;21(12):915-31. doi:10.1038/s41573-022-00552-x
8. Hingorani AD, Kuan V, Finan C, Kruger FA, Gaulton A, Chopade S, et al. Improving the odds of drug development success through human genomics: Modelling study. Sci Rep. 2019;9(1):18911. doi:10.1038/s41598-019-54849-w
9. Low LA, Mummery C, Berridge BR, Austin CP, Tagle DA. Organs-on-chips: Into the next decade. Nat Rev Drug Discov. 2021;20(5):345-61. doi:10.1038/s41573-020-0079-3
10. Munafò MR, Nosek BA, Bishop DVM, Button KS, Chambers CD, Percie du Sert N, et al. A manifesto for reproducible science. Nat Hum Behav. 2017;1(1):0021. doi:10.1038/s41562-016-0021
11. Errington TM, Mathur M, Soderberg CK, Denis A, Perfito N, Iorns E, et al. Investigating the replicability of preclinical cancer biology. Elife. 2021;10. doi:10.7554/eLife.71601
12. Voelkl B, Altman NS, Forsman A, Forstmeier W, Gurevitch J, Jaric I, et al. Reproducibility of animal research in light of biological variation. Nat Rev Neurosci. 2020;21(7):384-93. doi:10.1038/s41583-020-0313-3
13. King EA, Davis JW, Degner JF. Are drug targets with genetic support twice as likely to be approved? Revised estimates of the impact of genetic support for drug mechanisms on the probability of drug approval. PLoS Genet. 2019;15(12). doi:10.1371/journal.pgen.1008489
14. Minikel EV, Karczewski KJ, Martin HC, Cummings BB, Whiffin N, Rhodes D, et al. Evaluating drug targets through human loss-of-function genetic variation. Nature. 2020;581(7809):459-64. doi:10.1038/s41586-020-2267-z
15. Kim J, Koo BK, Knoblich JA. Human organoids: Model systems for human biology and medicine. Nat Rev Mol Cell Biol. 2020;21(10):571-84. doi:10.1038/s41580-020-0259-3
16. Burcham CL, Florence AJ, Johnson MD. Continuous manufacturing in pharmaceutical process development and manufacturing. Annu Rev Chem Biomol Eng. 2018;9:253-81. doi:10.1146/annurev-chembioeng-060817-084355
17. Nasr MM, Krumme M, Matsuda Y, Trout BL, Badman C, Mascia S, et al. Regulatory perspectives on continuous pharmaceutical manufacturing: Moving from theory to practice: September 26-27, 2016, International Symposium on the Continuous Manufacturing of Pharmaceuticals. J Pharm Sci. 2017;106(11):3199-206. doi:10.1016/j.xphs.2017.06.015
18. Yu LX, Kopcha M. The future of pharmaceutical quality and the path to get there. Int J Pharm. 2017;528(1-2):354-9. doi:10.1016/j.ijpharm.2017.06.039
19. Monticello TM, Jones TW, Dambach DM, Potter DM, Bolt MW, Liu M, et al. Current nonclinical testing paradigm enables safe entry to first-in-human clinical trials: The IQ Consortium nonclinical to clinical translational database. Toxicol Appl Pharmacol. 2017;334:100-9. doi:10.1016/j.taap.2017.09.006
20. Musuamba FT, Skottheim Rusten I, Lesage R, Russo G, Bursi R, Emili L, et al. Scientific and regulatory evaluation of mechanistic in silico drug and disease models in drug development: Building model credibility. CPT Pharmacometrics Syst Pharmacol. 2021;10(8):804-25. doi:10.1002/psp4.12669
21. Greenhalgh T, Wherton J, Papoutsi C, Lynch J, Hughes G, A'Court C, et al. Beyond adoption: A new framework for theorizing and evaluating nonadoption, abandonment, and challenges to the scale-up, spread, and sustainability of health and care technologies. J Med Internet Res. 2017;19(11). doi:10.2196/jmir.8775
22. Damschroder LJ, Reardon CM, Widerquist MAO, Lowery J. The updated Consolidated Framework for Implementation Research based on user feedback. Implement Sci. 2022;17(1):75. doi:10.1186/s13012-022-01245-0

23. Schuhmacher A, Wilisch L, Kuss M, Kandelbauer A, Hinder M, Gassmann O. R&D efficiency of leading pharmaceutical companies—a 20-year analysis. Drug Discov Today. 2021;26(8):1784-9. doi:10.1016/j.drudis.2021.05.005
24. Schweizer L, He J. Guiding principles of value creation through collaborative innovation in pharmaceutical research. Drug Discov Today. 2018;23(2):213-8. doi:10.1016/j.drudis.2017.09.003
25. Percie du Sert N, Hurst V, Ahluwalia A, Alam S, Avey MT, Baker M, et al. The ARRIVE guidelines 2.0: Updated guidelines for reporting animal research. PLoS Biol. 2020;18(7). doi:10.1371/journal.pbio.3000410
26. Kuemmel C, Yang Y, Zhang X, Florian J, Zhu H, Tegenge M, et al. Consideration of a credibility assessment framework in model-informed drug development: Potential application to physiologically based pharmacokinetic modeling and simulation. CPT Pharmacometrics Syst Pharmacol. 2020;9(1):21-8. doi:10.1002/psp4.12479
27. Häse F, Roch LM, Aspuru-Guzik A. Next-generation experimentation with self-driving laboratories. Trends Chem. 2019;1(3):282-91. doi:10.1016/j.trechm.2019.02.007
28. Austin CP. Opportunities and challenges in translational science. Clin Transl Sci. 2021;14(5):1629-47. doi:10.1111/cts.13055