



Formulation Models Should Predict Failure Modes Rather Than Only Optima When Designing Pharmaceutical Systems for Complex Natural Product Candidates

Marco De Luca^{1*}, Giulia Ferraro², Antonio Russo¹, Elena Rossi³

¹Department of Formulation Failure Mode Prediction, Faculty of Pharmacy, University of Naples Federico II, Naples, Italy.

²Department of Pharmaceutical Systems Design for Natural Products, Faculty of Pharmacy, University of Bologna, Bologna, Italy.

³Department of Optimized vs. Robust Formulation Design, Faculty of Pharmacy, University of Florence, Florence, Italy.

ABSTRACT

Machine-learning and data-driven formulation methods are increasingly used to predict pharmaceutical properties and identify compositions expected to maximize predefined performance objectives. For complex natural product candidates, however, a predicted optimum may be developmentally fragile if neighboring compositions, processing conditions, storage states, or physiological environments create physical instability, chemical degradation, manufacturing defects, precipitation, altered release, or loss of biological performance. This methodological perspective argues that formulation modeling should therefore represent failure explicitly rather than treating failure merely as the low-performing end of an optimization objective. We distinguish context-defined formulation failure from suboptimal performance and propose a failure-mode-first architecture in which desirable-performance targets remain analytically separate from typed physical, chemical, manufacturing, release, and biological failure targets. Failure probabilities are further distinguished from failure severity and prediction uncertainty. Stress testing is positioned as an active interrogation of predicted design space, while robust formulation regions are defined conceptually by acceptable performance maintained alongside bounded failure risk under prespecified relevant perturbations. Developmental updating should incorporate validated failures, near-failures, process changes, and distribution shifts rather than retaining only successful experiments. The framework does not establish universal failure thresholds, risk weights, or validated safe regions. Its purpose is instead to redirect formulation modeling from selecting isolated best predicted points toward identifying where performance is defensible, where models are uncertain, what can fail, and when data-driven prediction should yield to mechanistic investigation.

Key Words: Formulation modeling, Failure modes, Machine learning, Natural products, Robust design, Pharmaceutical development

eIJPPR 2026; 16(4):99-108

HOW TO CITE THIS ARTICLE: De Luca M, Ferraro G, Russo A, Rossi E. Formulation Models Should Predict Failure Modes Rather Than Only Optima When Designing Pharmaceutical Systems for Complex Natural Product Candidates. Int J Pharm Phytopharmacol Res. 2026;16(4):99-108. <https://doi.org/10.51847/izbWNZaggV>

INTRODUCTION

Data-driven methods now influence formulation selection, property prediction, process understanding, and experimental prioritization across pharmaceutical development [1]. Their attraction is straightforward:

formulation spaces can be combinatorially large, experimentation is costly, and relationships among drug properties, excipients, processing variables, and critical quality attributes are often nonlinear. A model that concentrates experimental effort around promising candidates can therefore be substantially more useful than

Corresponding author: Marco De Luca

Address: Department of Formulation Failure Mode Prediction, Faculty of Pharmacy, University of Naples Federico II, Naples, Italy.

E-mail: ✉ marco.deluca@unina.it

Received: 02 May 2026; **Revised:** 29 July 2026; **Accepted:** 07 August 2026

This is an **open access** journal, and articles are distributed under the terms of the Creative Commons Attribution-Non Commercial-ShareAlike 4.0 License, which allows others to remix, tweak, and build upon the work non-commercially, as long as appropriate credit is given and the new creations are licensed under the identical terms.



indiscriminate search. Yet the decision target matters as much as the algorithm. A model trained to identify the composition with the highest predicted dissolution, loading, stability score, or other desirable response is answering a different question from one trained to identify how that candidate can fail.

Contemporary formulation optimization increasingly combines surrogate modeling, Bayesian optimization, active learning, and multi-objective strategies to navigate complex design spaces [2]. Multi-objective optimization improves substantially on maximizing a single response because competing desirable properties can be represented simultaneously. Nevertheless, even a Pareto-efficient candidate can remain developmentally fragile if important failure events are absent from the objective set. A formulation may perform well under the condition used for model development yet undergo phase separation after hydration, degrade during processing, become difficult to manufacture after scale transfer, or lose its expected release behavior when composition or operating conditions move slightly away from the selected point.

This distinction is especially important for structurally complex natural product candidates. Their formulation problems may involve low or variable aqueous solubility, multiple ionizable or hydrogen-bonding functionalities, stereochemical complexity, susceptibility to chemical transformation, and formulation-dependent exposure. These features do not imply a common failure mechanism across natural products, but they increase the cost of assuming that a favorable modeled response is sufficient evidence of a robust pharmaceutical system.

The argument developed here is therefore methodological rather than algorithm-specific. A predicted optimum should be treated as incomplete unless the model also represents relevant ways in which the formulation can become unacceptable, the conditions under which those events emerge, and the uncertainty surrounding those predictions. Robustness, in this sense, is not maximum predicted performance. It is defensible performance within a region whose failure behavior has also been interrogated.

Why formulation modeling is biased toward best-case outcomes

Machine learning entered formulation development largely through prediction tasks in which known formulation or molecular descriptors were mapped to desirable product attributes and then used to support candidate ranking or experimental selection [3]. That orientation is productive, but it creates an asymmetry: successful formulations commonly generate structured measurements suitable for regression or optimization, whereas discarded, unstable, poorly processed, or incompletely characterized formulations may be sparsely measured, inconsistently labeled, or omitted from curated datasets. The resulting model can become highly informative about gradients toward desirable responses while remaining comparatively uninformed about the topology of failure.

The breadth of modern formulation platforms reinforces the distinction between coverage of formulation types and coverage of failure mechanisms. AI-enabled formulation systems can model several classes of pharmaceutical formulation and support prediction across heterogeneous design tasks [4]. Such breadth does not itself guarantee that physical instability, incompatibility, manufacturing failure, release transitions, or downstream biological loss have been encoded as separate outputs. Where the training objective rewards only desirable-response prediction, the optimizer can rationally select a point near a failure boundary that it was never asked to recognize.

The problem is therefore not optimization itself. It is an objective-definition problem. A best-performing formulation under measured conditions may be entirely appropriate when the response surface is smooth, uncertainty is low, and important failure mechanisms have been excluded experimentally. The concern arises when unacceptable events occupy sparse, discontinuous, stress-dependent, or poorly sampled regions. In those settings, maximizing an expected response can conceal the difference between a broad adequate region and a narrow peak surrounded by formulations that fail for different reasons.

Figure 1 summarizes the proposed shift: performance optimization remains part of formulation modeling, but failure taxonomy, failure-specific evidence, stress response, uncertainty, and developmental updating remain analytically visible rather than being collapsed into one optimum score.

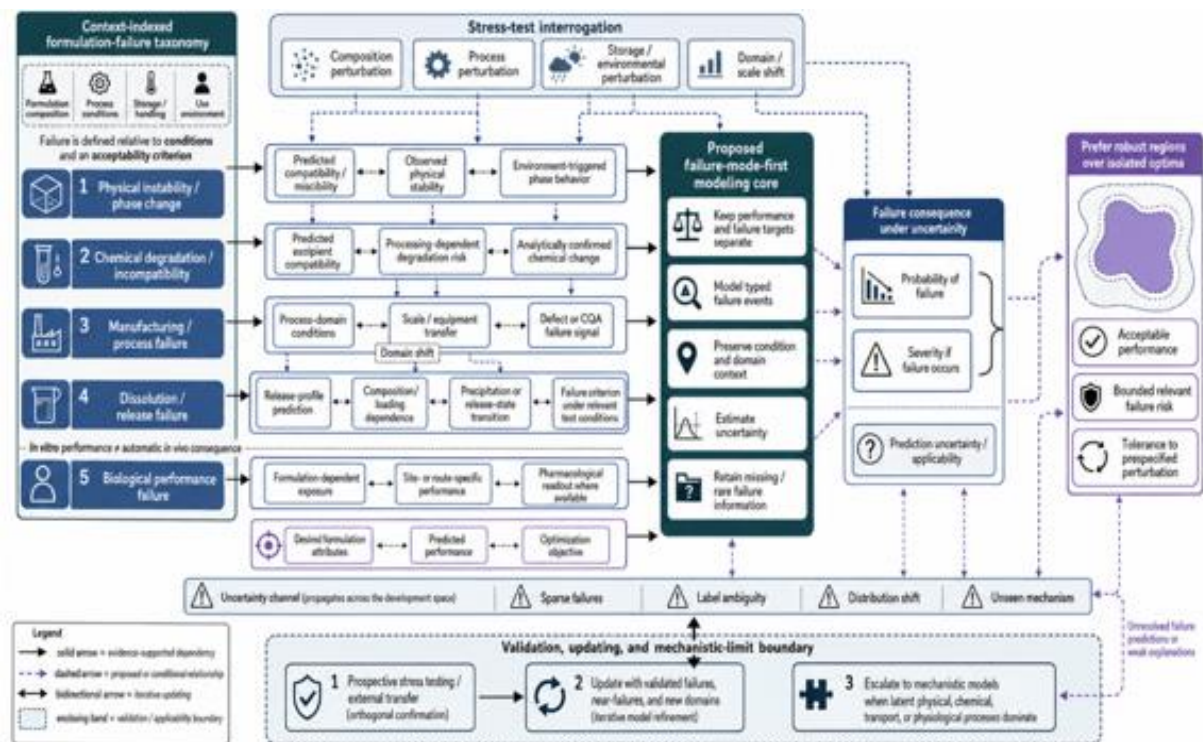


Figure 1. From optimum-seeking to failure-aware formulation modeling across a stress-conditioned development space

The distinct components of optimum-oriented formulation modeling and the information that a failure-aware strategy adds are compared in Table 1.

101.

Table 1. Method components that distinguish optimum-oriented from failure-aware formulation modeling

Method component	Conventional optimization emphasis	Failure-aware extension proposed here	Decision consequence
Target definition	Maximize or balance desirable responses	Preserve desirable responses and typed unacceptable events as distinct targets	A high-performing candidate can remain rejectable
Data retention	Prioritize well-characterized successful experiments	Retain validated failures, near-failures, and uncertain outcomes	Negative evidence becomes model information
Context	Formulation descriptors and measured conditions	Condition failure on composition, processing, storage, and use state	Failure is not treated as an intrinsic universal label
Model output	Predicted property or optimum	Performance, failure state, uncertainty, and applicability	Ranking becomes evidence-qualified
Design-space search	Identify best predicted point or Pareto set	Interrogate boundaries and nearby perturbations	Fragile optima can be distinguished from robust regions
Validation logic	Confirm expected performance	Test both expected performance and prespecified failure hypotheses	Validation can falsify an apparently favorable design

Note: The failure-aware column is a proposed methodological synthesis rather than a validated universal pharmaceutical-development standard.

What counts as a formulation failure?

A formulation failure should not be defined simply as “poor model performance” or as a formulation whose measured response is lower than the optimum. Those descriptions confound the behavior of the pharmaceutical system with the behavior of the prediction algorithm. The relevant event is instead a developmentally unacceptable state of the formulation, process, or resulting performance under specified conditions. Pharmaceutical-suspension work illustrates that stability can itself be treated as a supervised predictive outcome rather than only as a

covariate contributing to some other optimized response [5]. The broader methodological implication proposed here is that failure labels should be attached to scientifically interpretable events wherever the evidence permits. This definition has three consequences. First, failure is typed. Crystallization, phase separation, chemical degradation, unacceptable process defects, incomplete release, precipitation, and loss of biological performance are not interchangeable labels merely because each can terminate development. Second, failure is conditioned. A formulation may remain acceptable under one storage,



processing, hydration, dose, or use condition and fail under another. Third, failure is criterion-dependent. Whether a measured state is unacceptable depends on the development question and justified product-specific criteria; this perspective does not propose universal cut-offs.

Failure categories can also overlap. A physical transition can alter release; a process deviation can promote degradation; precipitation can change exposure. The purpose of typing failures is therefore not to create mutually exclusive bins. It is to prevent mechanistically different or decisionally different events from disappearing inside one scalar performance objective. A useful failure-aware dataset should consequently preserve the event observed, the conditions under which it occurred, the analytical level at which it was detected, and whether the evidence identifies mechanism or only association. This distinction becomes critical when physical instability is considered.

Physical instability and phase separation

Physical failure illustrates why a single “stability score” can be misleading. Machine-learning classification has been used to predict the physical stability of amorphous solid dispersions, while experimental phase-diagram work shows that hydration and composition can generate phase behavior that materially changes drug release, and thermodynamic calculations can separately estimate drug–polymer compatibility [6-8]. These bodies of evidence are related, but they do not measure the same thing. Predicted compatibility is not equivalent to observed storage stability, and neither automatically establishes how an apparently homogeneous system will behave after exposure to water or another environmental perturbation. A failure-aware model should therefore preserve at least three analytical states when the formulation requires them: thermodynamic or interaction propensity, experimentally observed physical state, and condition-triggered transformation. Combining them prematurely into one target risks training the model on labels whose meaning changes across experiments. Conversely, keeping them separate allows a prediction to be useful even when the mechanistic chain is incomplete: high estimated compatibility may reduce concern without proving stability, whereas experimentally observed phase separation constitutes a different level of evidence.

Stress testing follows naturally from this distinction. A candidate should not be characterized only at the nominal formulation and test condition if relevant perturbations are known to alter physical state. Composition, drug loading, humidity, hydration, temperature, and processing history may reveal boundaries invisible at the optimum. The appropriate perturbations are product- and mechanism-

dependent; the perspective therefore does not prescribe a universal stress panel. Its methodological claim is narrower: when a physical transition is decision-relevant, the model should represent the conditions that reveal it rather than average those conditions into a single apparent stability response. That requirement is particularly important when a physical state change subsequently alters dissolution, because the causal or temporal sequence should not disappear simply to simplify model architecture.

Chemical degradation and excipient interaction

Chemical failure requires an equally explicit separation between predicted incompatibility, degradation susceptibility, and experimentally demonstrated chemical transformation. Machine-learning methods can classify drug–excipient compatibility and thereby support preformulation screening [9]. Such predictions can be valuable for prioritizing experiments, but an incompatibility label does not identify a reaction pathway, establish degradant formation, or prove that the predicted interaction will occur under the processing and storage conditions of the intended product.

Direct formulation experiments show why the conditioning variables matter. In amorphous solid dispersions containing ritonavir, polymer choice altered the extent of thermal degradation during processing [10]. The permitted inference is not that a particular polymer class is universally degradative, but that chemical stability may depend on the joint drug–excipient–process context. Temperature history, residence time, local environment, water content, and formulation composition can therefore be relevant predictors or mechanistic variables when the underlying system warrants them.

For failure-aware modeling, chemical compatibility should consequently be used as one evidence layer rather than as a surrogate for the entire chemical-failure domain. A model may first estimate the probability of incompatibility, while analytical testing determines whether degradation actually occurs and mechanistic experiments clarify why. Those stages carry different evidential weight. Collapsing them can produce deceptively precise labels in which predictions, measured chemical change, and causal interpretation are treated as equivalent.

The proposed architecture instead records the event, its exposure or processing condition, the analytical evidence supporting it, and the degree of mechanistic resolution. This preserves a path from screening models to confirmatory analysis without requiring every formulation model to become a reaction-mechanism simulator. It also establishes the boundary for later development decisions: when chemically important behavior is controlled by latent reaction kinetics or states not represented in the training

data, purely empirical prediction may need to be supplemented or replaced by mechanistic modeling.

Manufacturing and processing failure

Manufacturing failure is not merely a formulation-property problem. A composition that performs adequately at laboratory scale can enter a different predictive domain when equipment, throughput, shear, thermal history, or control strategy changes. Transfer-learning work in pharmaceutical manufacturing, reviews of machine learning for lipid-nanoparticle formulation and processing, and deep-learning detection of tablet defects collectively show that scale transfer, process context, and product defects constitute distinct sources of development risk [11-13]. Their mechanisms differ, but each demonstrates why formulation descriptors alone may be insufficient.

A failure-aware model should therefore retain process domain as part of the prediction context. Scale transfer should trigger reassessment of applicability rather than automatic reuse of the original model. Likewise, a visible tablet defect can be predicted without establishing its causal root. The useful distinction is between predicting a failure phenotype and explaining why it arose. When the latter is required for corrective action, process knowledge or mechanistic investigation remains necessary.

Failure-aware process modeling should distinguish interpolation within a validated operating region from extrapolation to a new scale. Similar apparent product quality at two scales does not establish that identical input-output relationships govern both. Process transfer is therefore simultaneously a formulation question, a data-domain question, and, when unexpected failures emerge, a mechanism-discovery problem.

Dissolution and release failure

Dissolution models can use process and analytical information to provide rapid surrogate predictions, including real-time release-testing applications [14]. Process-state measurements during coating can likewise support prediction of dissolution profiles [15]. These capabilities are valuable, but a predicted profile is not itself a failure model. Failure exists only when the observed or predicted release behavior violates a justified condition-specific criterion.

Release is also unlikely to be uniformly smooth across formulation space. Drug loading, thermodynamic

descriptors, and manufacturing route can condition dissolution, while measurable in vitro differences do not necessarily translate proportionally into different in vivo performance [16-18]. A failure-first treatment should therefore preserve release transitions, precipitation or incomplete release when demonstrated, and the conditions that produce them rather than representing all observations by a single mean-response surface.

For natural-product candidates, this is especially relevant when supersaturation, precipitation, carrier disassembly, or matrix-controlled release makes an apparently favorable early dissolution response transient. Time and test condition should be preserved whenever they determine whether a performance gain persists.

This separation prevents two opposite errors: declaring every unfavorable dissolution shift a biological failure, or assuming that a favorable mean dissolution prediction excludes downstream risk. Release remains an important formulation endpoint, but its interpretation belongs to a level-specific chain of evidence.

Biological performance failure

For complex natural products, formulation can influence several levels simultaneously. A solid dispersion of andrographolide altered solubility, dissolution, pharmacokinetic behavior, and pharmacological readouts relative to comparator states [19]. Apigenin-loaded nanostructured lipid carriers similarly demonstrate that physicochemical carrier performance can be evaluated alongside biological outcomes in a distinct delivery setting [20]. These examples show that formulation success cannot always be judged at the dissolution level alone.

They do not establish a common natural-product failure mechanism. Biological performance depends on candidate chemistry, route, carrier, exposure, experimental model, and the relevance of the measured endpoint. Accordingly, the proposed framework treats biological failure as a separate evidence level. Formulation-dependent exposure or pharmacological response may inform development, but neither should be inferred automatically from an upstream physical or release measurement.

The stepwise evidence logic that should be preserved before a formulation is labeled as failed is summarized in **Table 2**.

Table 2. Evidence-preserving workflow from formulation signal to developmentally interpretable failure

Step	Evidence to preserve	Decision question	Interpretation boundary
Detect	Measured or predicted abnormal state	What changed?	Signal is not yet mechanism
Condition	Composition, process, storage, or use state	Under what conditions?	Failure is context dependent
Type	Physical, chemical, manufacturing, release, or biological event	What kind of failure is implicated?	Categories can coexist

Confirm	Orthogonal or direct measurement where required	Is the event reproducible and analytically supported?	Prediction is not confirmation
Interpret	Mechanistic evidence and downstream consequence	Why does it matter for development?	In vitro change is not automatically biological failure

Note: The workflow is a proposed organizational device; required measurements and acceptance criteria remain product- and mechanism-specific.

Figure 2 shows how those evidence levels can remain separate while still contributing to one formulation decision architecture.

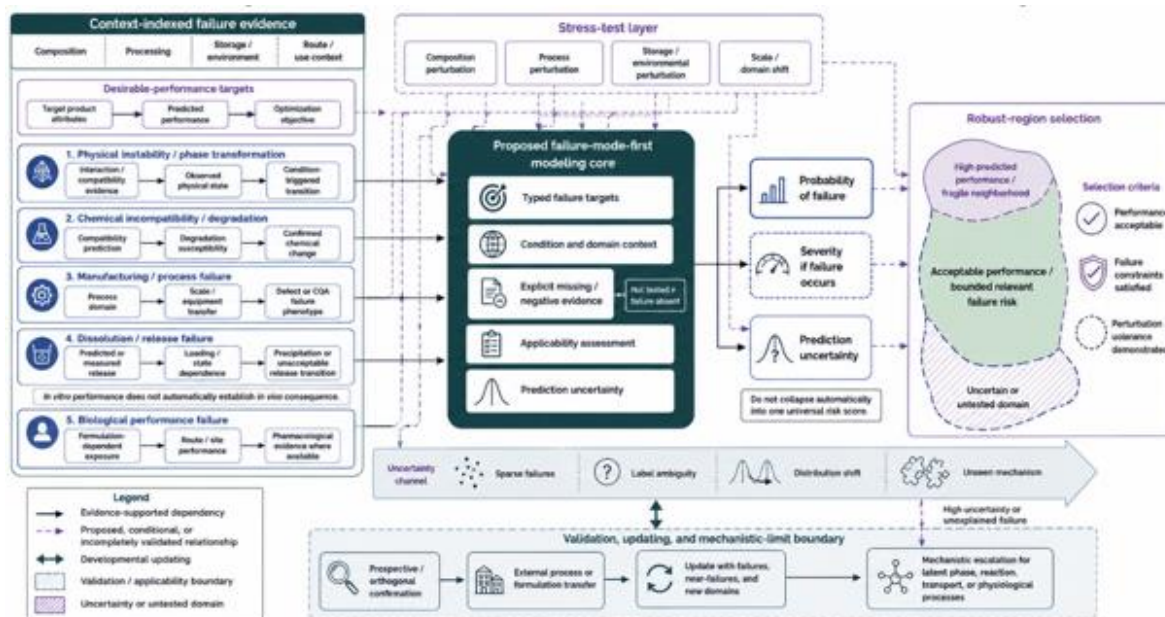


Figure 2. Layered failure evidence architecture for robust formulation decisions under uncertainty

A failure-mode-first modeling architecture

Sparse experimental datasets make exhaustive mapping of formulation space unrealistic; machine learning can nevertheless assist formulation design when experiments are limited [21]. The proposed architecture therefore does not require a complete catalogue of every possible failure. It requires that known or scientifically plausible failure modes be represented explicitly enough that absence of a failure label is not mistaken for evidence of acceptability. Rare failures create a second problem: a classifier can appear successful while performing poorly on the minority class. Threshold-adjustment work for imbalanced chemical datasets demonstrates why default classification conventions may be inappropriate when the event of interest is uncommon [22]. Distribution shift adds another layer because uncertainty estimates themselves can change when the prediction domain moves away from training data [23].

The architecture should also avoid treating every absence of an observed failure as a negative class. An experiment that was never stressed or never assayed for a specific degradation or biological endpoint is epistemically different from one in which that failure was actively sought

and not observed. Preserving this distinction reduces false confidence created by experimental missingness.

The proposed modeling core consequently contains separate performance targets, typed failure targets, applicability information, and uncertainty. Missing failure information should remain missing rather than silently recoded as success. Outputs are interpreted at a decision layer, not collapsed automatically into one universal score. The architecture is deliberately model-agnostic: classification, regression, survival-type formulations, ensembles, or hybrid mechanistic/data-driven models may be appropriate depending on the failure process.

Stress testing the predicted design space

Failure models become useful only if their boundaries are challenged. Active-learning formulation workflows show that experiments can be selected iteratively rather than uniformly across a design space [24]. Generative formulation approaches similarly expand the set of candidate designs but still require empirical verification [25]. The proposed extension is to direct part of that experimental budget toward predicted boundaries, uncertainty regions, and scientifically plausible failure conditions.

Stress testing should perturb variables that can expose relevant failure: composition, processing history, storage environment, scale, or use condition. The aim is not to create artificially harsh tests or a universal challenge panel. It is to ask whether the candidate remains acceptable under prespecified perturbations meaningful to its development context.

Boundary-directed experiments should include both expected failure regions and high-uncertainty regions because the two are not synonymous. A model may predict low failure probability with high uncertainty, or high failure probability confidently; those states justify different experimental priorities. A model that survives only at the nominal optimum has supplied a location, not evidence of robustness.

Modeling probability and severity of failure together

Release-prediction models for polymeric long-acting injectables demonstrate how continuous formulation behavior can guide candidate design [26]. Failure-aware decision making adds a separate question: even when the probability of an undesirable event can be estimated, how consequential would that event be?

Probability and severity should therefore remain distinct. A frequent but readily correctable process deviation need not carry the same development consequence as a rare event causing irreversible chemical instability. Severity is not a property that the approved evidence supplies universally; it must be defined for the product, development stage, and decision. Prediction uncertainty is separate again. High uncertainty can justify additional

testing without implying that the underlying failure itself is severe.

This perspective therefore rejects an automatic universal probability-times-severity formula. Where quantitative utilities are justified, they should be prespecified and validated for the relevant decision. Otherwise, a structured qualitative separation is preferable to fabricated precision.

Choosing robust regions rather than single optima

Closed-loop active learning can navigate formulation composition toward desired release behavior through sequential experimentation [27]. The methodological extension proposed here is to replace the question “Which point is best?” with “Which connected region remains acceptable when both performance and relevant failure constraints are considered?”

A robust region need not contain the numerically highest predicted response. It may instead contain formulations with slightly lower nominal performance but greater tolerance to composition, process, storage, or domain perturbation. Such a region is meaningful only relative to specified constraints; it is not a universal safe zone.

Connectedness also matters conceptually. Two isolated acceptable points separated by a region of process or stability failure are not equivalent to one contiguous region tolerant to ordinary manufacturing variability. The framework therefore treats design-space geometry qualitatively unless sufficient data justify formal estimation.

Table 3 contrasts this decision logic with common optimization approaches.

Table 3. How failure-first robust-region selection differs from common formulation optimization strategies

Approach	Primary decision target	Treatment of failure	Uncertainty and stress	Main limitation for robust selection
Single-response optimization	Best predicted response	Usually implicit in low performance	Often secondary	Can favor a narrow fragile optimum
Multi-objective optimization	Trade-off among desired responses	Explicit only if encoded as objective or constraint	Method dependent	Pareto efficiency does not guarantee failure coverage
Active learning / Bayesian search	Informative or promising next experiment	Depends on acquisition objective	Can target uncertainty	May still concentrate on performance gain
Proposed failure-first selection	Acceptable region satisfying performance and typed failure constraints	Explicit and mode-specific	Integral to selection and re-testing	Requires reliable failure definitions and sufficient negative evidence

Note: The comparison is methodological; no row implies universal superiority across formulation problems.

Updating failure models during pharmaceutical development

A formulation model should not be frozen when the development system changes. Iterative formulation experimentation provides a practical basis for sequential updating [24]. Scale transfer likewise creates a reason to reassess model applicability rather than assuming stationarity [11]. Distribution shift may appear through

new chemistry, process conditions, equipment, analytical methods, or biological contexts and can degrade confidence in otherwise accurate models [23].

Updating should retain validated failures, near-failures, and informative negative experiments after appropriate quality adjudication. It should also distinguish a change in failure prevalence from emergence of a genuinely new failure mechanism. The former may be addressable by

recalibration or new data; the latter may indicate that the representation itself is incomplete.

Model revision should be versioned against the experimental domain that justified it, so improved fit after new data is not confused with preserved validity in earlier or external domains. Developmental updating is thus both statistical and scientific: the model changes, but so can the definition of what must be modeled.

Limitations and cases requiring mechanistic models

Data-driven failure prediction cannot replace mechanistic understanding when critical state variables are absent from the training representation. Integrated AI and physiologically based pharmacokinetic modeling for amorphous solid dispersions illustrates one route for linking formulation-level prediction with mechanistic description of downstream fate [28]. The broader escalation principle proposed here is conditional: mechanistic models become especially important when failure depends on phase transitions, reaction kinetics, transport, physiology, or other latent processes that empirical correlations cannot reliably extrapolate.

The framework also remains limited by sparse failures, inconsistent labels, uncertain severity definitions, distribution shift, and unobserved mechanisms. Class-imbalance methods can improve decision handling [22], but they cannot manufacture missing scientific evidence. Uncertainty estimation can flag unreliable regions [23], but uncertainty is not proof of failure or acceptability.

Mechanistic escalation should not be romanticized as certainty. Thermodynamic, kinetic, transport, or PBPK models contain assumptions and parameter uncertainty; their value is that they can represent scientifically relevant state variables that a purely empirical representation may omit. No universal failure threshold, stress protocol, robust-region geometry, or transfer criterion is established here.

CONCLUSION

Formulation modeling should not stop after identifying the composition with the most attractive predicted response. For complex natural product candidates in particular, the development question is broader: what can fail, under which conditions, how consequential might that failure be, how certain is the prediction, and does acceptable performance persist around the selected formulation?

The failure-mode-first perspective proposed here preserves physical, chemical, manufacturing, release, and biological failures as distinct but interacting evidence domains. It couples them to stress testing, uncertainty, robust-region selection, developmental updating, and mechanistic escalation without claiming a universal risk algorithm. Its

practical value lies in changing what a formulation model is asked to explain. Rather than treating unsuccessful observations as inconvenient departures from an optimum, they become information about the boundaries of the pharmaceutical system.

The resulting objective is not the mathematically best point in isolation. It is a formulation region whose performance remains defensible because relevant failure modes have been sought, represented, challenged, and interpreted within their evidential limits.

Acknowledgments: None

Conflict of interest: None

Financial support: None

Ethics statement: None

REFERENCES

- [1] Bao Z, Bufton J, Hickman RJ, Aspuru-Guzik A, Bannigan P, Allen C. Revolutionizing drug formulation development: The increasing impact of machine learning. *Adv Drug Deliv Rev.* 2023;202:115108. doi:10.1016/j.addr.2023.115108
- [2] Ros H, Chan N, Cook MT, Shorthouse D. Artificial intelligence and machine learning guided optimization in drug delivery. *Adv Drug Deliv Rev.* 2026;232:115781. doi:10.1016/j.addr.2026.115781
- [3] Bannigan P, Aldeghi M, Bao Z, Häse F, Aspuru-Guzik A, Allen C. Machine learning directed drug formulation development. *Adv Drug Deliv Rev.* 2021;175:113806. doi:10.1016/j.addr.2021.05.016
- [4] Dong J, Wu Z, Xu H, Ouyang D. FormulationAI: A novel web-based platform for drug formulation design driven by artificial intelligence. *Brief Bioinform.* 2024;25(1):bbad419. doi:10.1093/bib/bbad419
- [5] Zulfeari N, Wang F, Mustafaova SS, Parhizkar M, Holm R. Machine learning strengthened formulation design of pharmaceutical suspensions. *Int J Pharm.* 2025;668:124967. doi:10.1016/j.ijpharm.2024.124967
- [6] Han R, Xiong H, Ye Z, Yang Y, Huang T, Jing Q, et al. Predicting physical stability of solid dispersions by machine learning techniques. *J Control Release.* 2019;311-312:16-25. doi:10.1016/j.jconrel.2019.08.030
- [7] Deac A, Luebbert C, Qi Q, Courtney RM, Indulkar AS, Gao Y, et al. Dissolution mechanisms of amorphous solid dispersions: Application of ternary phase diagrams to explain release behavior. *Mol*

- Pharm. 2024;21(4):1900-18.
 doi:10.1021/acs.molpharmaceut.3c01179
- [8] Antolović I, Vrabec J, Klajmon M. COSMOPharm: Drug-polymer compatibility of pharmaceutical amorphous solid dispersions from COSMO-SAC. *Mol Pharm.* 2024;21(9):4395-415.
 doi:10.1021/acs.molpharmaceut.4c00342
- [9] Hang NT, Long NT, Duy ND, Chien NN, Van Phuong N. Towards safer and efficient formulations: Machine learning approaches to predict drug-excipient compatibility. *Int J Pharm.* 2024;653:123884.
 doi:10.1016/j.ijpharm.2024.123884
- [10] de Alvarenga BR Jr, Moseson DE, Carneiro RL, Taylor LS. Impact of polymer type on thermal degradation of amorphous solid dispersions containing ritonavir. *Mol Pharm.* 2022;19(1):332-44.
 doi:10.1021/acs.molpharmaceut.1c00823
- [11] Yaginuma K, Matsunami K, Descamps L, Ryckaert A, De Beer T. Hybrid modeling of T-shaped partial least squares regression and transfer learning for formulation and manufacturing process development of new drug products. *Int J Pharm.* 2024;662:124463.
 doi:10.1016/j.ijpharm.2024.124463
- [12] Dorsey PJ, Lau CL, Chang TC, Doerschuk PC, D'Addio SM. Review of machine learning for lipid nanoparticle formulation and process development. *J Pharm Sci.* 2024;113(12):3413-33.
 doi:10.1016/j.xphs.2024.09.015
- [13] Diószegi A, Ficzer M, Mészáros LA, Péterfi O, Farkas A, Galata DL, et al. Automated tablet defect detection and the prediction of disintegration time and crushing strength with deep learning based on tablet surface images. *Int J Pharm.* 2024;667(Pt A):124896.
 doi:10.1016/j.ijpharm.2024.124896
- [14] Galata DL, Könyves Z, Nagy B, Novák M, Mészáros LA, Szabó E, et al. Real-time release testing of dissolution based on surrogate models developed by machine learning algorithms using NIR spectra, compression force and particle size distribution as input data. *Int J Pharm.* 2021;597:120338.
 doi:10.1016/j.ijpharm.2021.120338
- [15] Sacher S, Fink E, Alva C, Afonso Ulrich JA, Doğan A, Herndler V, et al. Real-time prediction of dissolution profiles of coated oral dosage forms. *Int J Pharm.* 2024;666:124841.
 doi:10.1016/j.ijpharm.2024.124841
- [16] Saboo S, Moseson DE, Kestur US, Taylor LS. Patterns of drug release as a function of drug loading from amorphous solid dispersions: A comparison of five different polymers. *Eur J Pharm Sci.* 2020;155:105514. doi:10.1016/j.ejps.2020.105514
- [17] Ge K, Shen J, Chen H, Wang J, Cui Y, Shen C, et al. Advancing amorphous solid dispersions design: Insights into dissolution kinetics via thermodynamic descriptor and machine learning. *Mol Pharm.* 2025;22(7):4259-68.
 doi:10.1021/acs.molpharmaceut.5c00518
- [18] Trenkenschuh E, Blattner SM, Hirsh D, Hoffmann R, Luebbert C, Schaefer K. Development of ternary amorphous solid dispersions manufactured by hot-melt extrusion and spray-drying—comparison of in vitro and in vivo performance. *Mol Pharm.* 2024;21(3):1309-20.
 doi:10.1021/acs.molpharmaceut.3c00696
- [19] Haq SA, Paudwal G, Banjare N, Andrabi NI, Wazir P, Nandi U, et al. Sustained release polymer and surfactant based solid dispersion of andrographolide exhibited improved solubility, dissolution, pharmacokinetics, and pharmacological activity. *Int J Pharm.* 2024;651:123786.
 doi:10.1016/j.ijpharm.2024.123786
- [20] Bonilla-Vidal L, Espina M, García ML, Baldomà L, Badia J, González JA, et al. Novel nanostructured lipid carriers loading Apigenin for anterior segment ocular pathologies. *Int J Pharm.* 2024;658:124222.
 doi:10.1016/j.ijpharm.2024.124222
- [21] Narayanan H, Dingfelder F, Condado Morales I, Patel B, Heding KE, Bjelke JR, et al. Design of biopharmaceutical formulations accelerated by machine learning. *Mol Pharm.* 2021;18(10):3843-53.
 doi:10.1021/acs.molpharmaceut.1c00469
- [22] Esposito C, Landrum GA, Schneider N, Stiefl N, Riniker S. GHOST: Adjusting the decision threshold to handle imbalanced data in machine learning. *J Chem Inf Model.* 2021;61(6):2623-40.
 doi:10.1021/acs.jcim.1c00160
- [23] Parrondo-Pizarro R, Lanini J, Rodríguez-Pérez R. Uncertainty quantification in molecular machine learning for property predictions under data shifts. *J Chem Inf Model.* 2026;66(2):923-35.
 doi:10.1021/acs.jcim.5c02381
- [24] Bao Z, Le Dévédec F, Huynh S. A data-driven workflow for nanomedicine optimization using active learning and automated experimentation. *Mol Pharm.* 2025;22(12):7465-77.
 doi:10.1021/acs.molpharmaceut.5c00933
- [25] Hornick T, Mao C, Koynov A, Yawman P, Thool P, Salish K, et al. In silico formulation optimization and particle engineering of pharmaceutical products using a generative artificial intelligence structure synthesis method. *Nat Commun.* 2024;15(1):9622.
 doi:10.1038/s41467-024-54011-9
- [26] Bannigan P, Bao Z, Hickman RJ, Aldeghi M, Häse F, Aspuru-Guzik A, et al. Machine learning models to accelerate the design of polymeric long-acting

- injectables. Nat Commun. 2023;14(1):35. doi:10.1038/s41467-022-35343-w
- [27] Cheong E, Radford DC, Gormley AJ. Automated active learning to optimize hydrogel drug release profiles. J Control Release. 2026;391:114602. doi:10.1016/j.jconrel.2026.114602
- [28] Zhu J, Xiong P, Wang W, Lu T, Ouyang D. Integrating artificial intelligence and physiologically based pharmacokinetic modeling to predict in vitro and in vivo fate of amorphous solid dispersions. J Control Release. 2025;386:114123. doi:10.1016/j.jconrel.2025.114123