Pharmacophore

ISSN-2229-5402

Journal home page: http://www.pharmacophorejournal.com

When Learned Dose Predictions Conflict with Pharmacological Knowledge: A Theoretical Framework for Reconciling Artificial Intelligence with Mechanistic Pharmacokinetics, Pharmacodynamics, and Therapeutic Drug Monitoring

Peter Novak[1]*, Leila Haddad[2], Johan Nilsson[3], Aisha Bello[4]

1. ***Department of Clinical Pharmacology and AI-Based Dosing, Charles University, Prague, Czech Republic.***
2. ***Department of Pharmacokinetics and Therapeutic Drug Monitoring, American University of Beirut, Beirut, Lebanon.***
3. ***Department of Clinical Pharmacy and Pharmacodynamics, Uppsala University, Uppsala, Sweden.***
4. ***Department of Pharmaceutical Care and Precision Dosing, University of Lagos, Lagos, Nigeria.***

ARTICLE INFO

Received:
29 May 2026
Received in revised form:
01 October 2026
Accepted:
04 October 2026
Available online:
28 October 2026

Keywords: Machine learning, Model-informed precision dosing, Pharmacokinetics, Pharmacodynamics, Therapeutic drug monitoring, Dose reconciliation

ABSTRACT

Artificial intelligence is increasingly capable of generating patient-specific predictions of drug exposure, response, or dose requirements, while pharmacokinetic and pharmacodynamic models provide mechanistically structured estimates derived from drug disposition, exposure–response relationships, prior population information, and patient covariates. These approaches need not converge. Their disagreement creates a clinically important problem because superior retrospective prediction does not necessarily identify the more credible dose for an individual patient, whereas mechanistic plausibility does not guarantee that a model remains applicable when physiology, treatment context, or population characteristics depart from its assumptions. This theoretical reconciliation article examines how such conflicts should be interpreted when learned predictions, mechanistic PK/PD expectations, and therapeutic drug monitoring provide partially overlapping but non-equivalent information. The central argument is that no modality should receive automatic priority. Disagreement should first be diagnosed according to its likely source, including covariate representation, training-support limitations, model-form uncertainty, population transportability, changing patient biology, target definition, or measurement quality. Reconciliation should then depend on the provenance, clinical relevance, temporal proximity, and validation status of the competing evidence. Therapeutic drug monitoring becomes particularly important when valid patient-linked observations can discriminate among previously plausible predictions, but its evidentiary weight depends on sampling informativeness, analytical validity, and the exposure or response target being estimated. The resulting framework treats concordant, discordant, and unresolved states as provisional clinical states subject to monitoring and revision. Dose selection therefore becomes an evidence-updating process rather than a contest between artificial intelligence and pharmacological mechanism.

*This is an **open-access** article distributed under the terms of the Creative Commons Attribution-Non Commercial-Share Alike 4.0 License, which allows others to remix, and build upon the work non commercially.*

To Cite This Article: Novak P, Haddad L, Nilsson J, Bello A. When Learned Dose Predictions Conflict with Pharmacological Knowledge: A Theoretical Framework for Reconciling Artificial Intelligence with Mechanistic Pharmacokinetics, Pharmacodynamics, and Therapeutic Drug Monitoring. Pharmacophore. 2026;17(5):103-11. https://doi.org/10.51847/8YU49A2ne4

Introduction

Machine-learning methods are increasingly used to estimate drug concentrations, pharmacokinetic quantities, and patient-specific exposure trajectories from combinations of demographic, laboratory, dosing, molecular, and longitudinal clinical information. Reviews of this literature show that learned predictors can achieve useful concentration-prediction performance and, in some settings, performance comparable with established population-pharmacokinetic approaches [1]. Machine learning has also moved beyond concentration estimation toward explicit dose individualization, including a priori prediction before

Corresponding Author: Peter Novak; Department of Clinical Pharmacology and AI-Based Dosing, Charles University, Prague, Czech Republic. E-mail: peter.novak@cuni.cz.

treatment and a posteriori adaptation after treatment information becomes available [2]. The practical implication is that clinicians and clinical pharmacologists may increasingly encounter two technically credible but numerically different recommendations: one derived mainly from learned associations and another from mechanistic pharmacological reasoning.

Mechanistic dosing should not be treated as an error-free reference against which such predictions are automatically judged. Model-informed precision dosing combines population models, individual covariates, drug exposure information, pharmacodynamic targets, biomarkers, and repeated patient observations to refine treatment decisions [3]. Its validity depends on model qualification, transportability, parameter uncertainty, appropriate exposure or effect targets, and the quality of the data entering the model [4]. Moving those calculations to the bedside introduces additional problems involving workflow, software, model selection, timeliness of information, and the clinical interpretation of uncertainty [5]. A mechanistic estimate is therefore better understood as a structured pharmacological prior whose credibility depends on how well its assumptions represent the current patient and decision.

Implementation evidence reinforces the difference between predictive performance and clinically trustworthy dosing. A systematic review of model-informed precision-dosing implementation identified barriers and facilitators spanning model development, validation, technical infrastructure, organizational processes, user knowledge, regulation, and clinical workflow [6]. Predictive artificial intelligence faces a parallel problem. Recent work in clinical pharmacology has argued that heterogeneous datasets, tasks, evaluation metrics, and benchmarking practices make superficially favorable model comparisons difficult to translate into reliable claims about clinical applicability [7]. A model may fit historical observations well while remaining poorly calibrated for an unusual patient, a changing physiological state, or a treatment decision different from the one represented in its training data.

The relevant question is consequently not whether artificial intelligence or PK/PD modeling is intrinsically superior. It is how to interpret disagreement between them when each may be correct within a different evidentiary domain, partially misspecified, or insufficiently individualized. The problem becomes especially important once therapeutic drug monitoring generates patient-specific observations that were unavailable when an initial dose was chosen. The analytical task of this article is therefore to identify why learned and mechanistic dose estimates diverge, determine which evidence can discriminate among competing explanations, and establish how dose selection should remain revisable when disagreement cannot yet be resolved.

Sources of Conflict between Learned and Mechanistic Dose Predictions

A first source of conflict is representational. Learned systems can incorporate covariates and nonlinear interactions that are absent from a conventional population PK model, whereas mechanistic models preserve structural relationships that a predictive algorithm may capture only indirectly. In critically ill patients receiving vancomycin, an interpretable hybrid workflow showed that machine learning could identify relationships between clinical features and individual pharmacokinetic parameters that were not represented by traditional covariate structures [8]. A separate iohexol-clearance application demonstrated that machine learning can complement population-pharmacokinetic and Bayesian estimation rather than merely compete with them [9]. Comparative work on human intravenous pharmacokinetic profiles likewise shows that data-driven, mechanistic, and hybrid approaches encode different information and therefore need not generate identical concentration–time predictions [10]. Differences between machine-learning architectures can themselves produce different PK predictions even when the broad task is unchanged [11]. The precision-dosing literature in oncology further shows substantial heterogeneity in algorithms, inputs, outputs, validation strategies, and intended decisions [12]. Disagreement can therefore arise before any one model is demonstrably wrong.

This representational difference becomes clinically important when a patient lies near the boundary of what a model has learned or what a mechanistic structure assumes. A learned predictor may be well calibrated inside a densely represented training region yet extrapolate poorly when renal function, interacting therapy, organ support, body composition, adherence, or disease state differs from the development population. A mechanistic model may extrapolate more coherently under some of those conditions, but only if its structural assumptions and parameter relationships remain valid. Conversely, an established PK model may systematically omit patient characteristics that a newer learned model detects as informative. The direction of disagreement therefore cannot be used as a shortcut for identifying the more credible estimate. What matters is the reason that the predictions have separated.

A second source of conflict is uncertainty among mechanistic models themselves. External evaluations show that several population PK models for the same drug can yield different predictive performance when transported into a new population [13]. Large-scale vancomycin work has further shown that patient characteristics can be used to rank or average several candidate PK models and that such strategies can outperform reliance on a single prespecified model for concentration prediction [14]. Simulation work on initial amoxicillin dosing illustrates the same problem before therapeutic drug-monitoring information becomes available: the most appropriate model may depend on how closely the patient resembles the population from which candidate models were derived [15]. For protein-kinase inhibitors, model averaging has also provided a means of reducing sensitivity to any one prior model under specified conditions [16]. These findings matter because they undermine a simple AI-versus-mechanism dichotomy. A learned prediction may disagree not with “pharmacology” in the abstract but with one of several plausible mechanistic parameterizations.

Hybrid modeling provides evidence that the two traditions are potentially complementary, but hybridization does not abolish the problem of adjudication. Current recommendations for pharmacometric–machine-learning models emphasize the need for transparent model construction, benchmarking, uncertainty characterization, external validation, and explicit treatment of error

propagation [17]. Reviews of hybrid mechanistic–machine-learning PK/PD approaches similarly identify opportunities to combine biological structure with flexible data-driven components while retaining concerns about missingness, data leakage, device or measurement drift, identifiability, and domain transport [18]. Neural PK/PD modeling has demonstrated that learned differential dynamics can be embedded within recognizable pharmacological structures and used to predict longitudinal response patterns [19]. In oncology, mechanistically derived PK/PD information has also been used as input to machine-learning systems for predicting chemotherapy-induced neutropenia [20]. These examples show that reconciliation can occur inside a model architecture, but they do not justify assuming that every hybrid model inherits the strengths of both parent approaches without also inheriting their uncertainties.

The practical implication is that prediction conflict should be treated as a diagnostic signal rather than an immediate instruction to select a winner. A disagreement is informative only when it can be linked to competing explanations and to observations capable of discriminating among them. Apparent conflict may reflect differences in covariate representation, model structure, population relevance, intended estimand, sampling time, exposure target, or current patient state. It may also be an artefact of comparing predictions that are not genuinely commensurable. The analytical distinction is therefore between a conflict signature—the observable pattern of disagreement—and a conflict source—the underlying reason that two otherwise credible approaches have separated. The relationship between those sources, the evidence needed to investigate them, and the shortcuts that should be avoided is summarized in **Table 1**.

Table 1. Diagnosing disagreement between learned dose predictions and mechanistic PK/PD expectations

Conflict unit	Typical diagnostic signature	Plausible source of disagreement	Evidence that can discriminate	Inappropriate shortcut
Covariate-support mismatch	Learned prediction changes sharply for a patient with uncommon clinical characteristics while the mechanistic estimate remains comparatively stable	Patient lies outside, or near the edge of, the learned training distribution	Training-support assessment, external calibration, clinically similar validation cases, mechanistic plausibility	Assuming high overall model accuracy applies equally to the atypical patient
Mechanistic structural mismatch	Mechanistic prediction remains internally coherent but conflicts with repeated observations or clinically plausible patient-specific patterns	Model structure or covariate relationships inadequately represent current physiology	Residual patterns, alternative PK structures, patient-specific physiology, repeated concentration data	Treating pharmacological structure as proof that the model is correctly specified
Competing mechanistic priors	Different credible population models produce materially different initial doses or exposures	Population transportability, parameter priors, model-development differences	External validation, population similarity, model averaging or relevance assessment	Selecting the most familiar published model without target-population evaluation
Target or estimand mismatch	Two systems appear discordant despite predicting different exposure windows, effect targets, or clinical quantities	Predictions are not directly comparable	Alignment of prediction horizon, endpoint, dose objective and uncertainty representation	Averaging quantities that do not estimate the same decision target
Temporal patient-state shift	Earlier prediction becomes less consistent with newly emerging physiology or treatment response	Renal, hepatic, inflammatory, interaction, adherence, organ-support, or disease-state change	Updated clinical measurements, repeated PK observations, revised covariates	Preserving the original prediction because it was well validated at baseline
Shared or correlated error	Learned and mechanistic estimates agree despite later evidence contradicting both	Common training data, shared assumptions, common covariate errors or incorrect target	Independent validation evidence, orthogonal clinical measurements, subsequent TDM	Treating agreement itself as validation
Hybrid-model conflict	Integrated model improves average prediction but component contributions or uncertainty remain unclear	Learned correction compensates for mechanistic misspecification, or vice versa	Component-level validation, ablation/benchmarking, external testing, uncertainty analysis	Assuming hybrid architecture automatically resolves causal or clinical uncertainty

The principal sources of disagreement between learned and mechanistic dose predictions, together with the corresponding evidence required to discriminate among competing explanations, are illustrated in **Figure 1**.

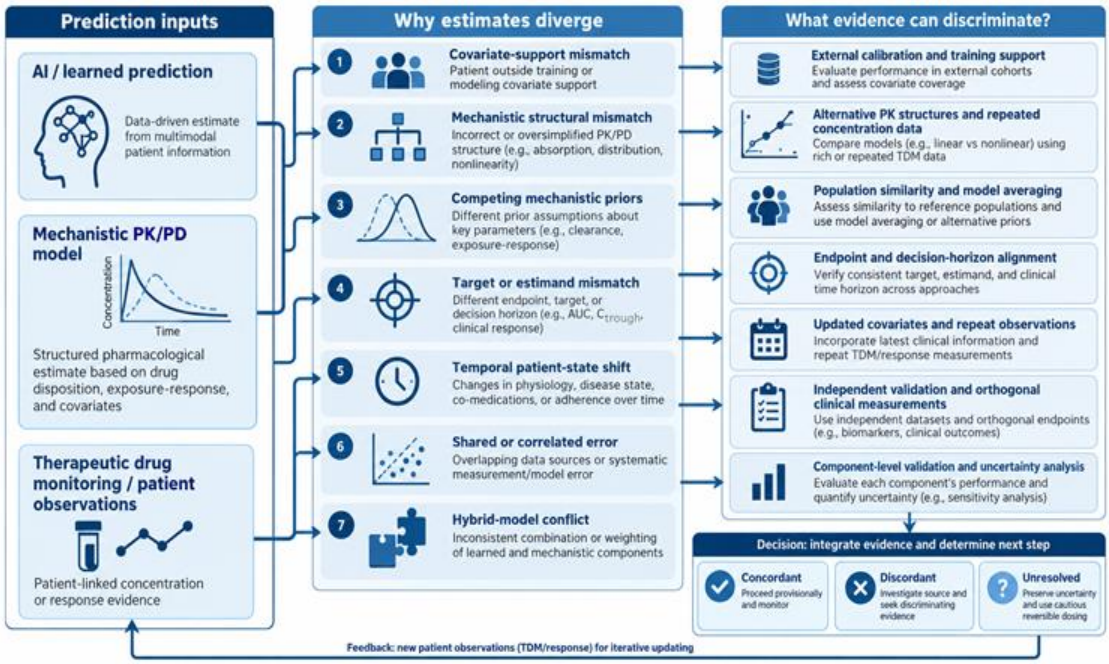


Figure 1. Diagnostic pathways for identifying and resolving disagreement between learned and mechanistic dose predictions.

Reconciling AI, PK/PD and Therapeutic Drug Monitoring

Therapeutic drug monitoring changes the epistemic structure of the dose decision because it introduces patient-linked observations after treatment has begun. In antibiotics, the progression from conventional TDM toward model-informed precision dosing has been driven by the ability to combine measured concentrations with population PK information and pharmacodynamic targets rather than interpret a concentration in isolation [21]. Digital-health approaches extend this idea by treating dosing as a longitudinal process in which laboratory values, physiological changes, adherence information, treatment events, and repeated measurements can alter the relevance of earlier predictions [22]. Newer evidence should not automatically override earlier evidence, but an observation obtained from the patient receiving the drug can acquire greater decision relevance when it is analytically reliable and informative about the competing models.

The updating process also depends on the infrastructure through which evidence reaches the clinician. Evaluations of model-informed precision-dosing software show substantial variation in model availability, validation, electronic-record integration, usability, reporting, and decision-support functionality [23]. Methodological work on continued-learning approaches further demonstrates that pharmacometric models can be updated as new clinical data accumulate, turning model development into a potentially iterative process rather than a permanently frozen predeployment event [24]. That flexibility creates an important safeguard only when updates remain traceable and validated. Uncontrolled adaptation could otherwise transform local practice patterns, measurement error, or selection bias into apparent evidence that the model has improved.

Therapeutic drug monitoring itself must therefore be qualified before it is allowed to adjudicate a disagreement. Machine-learning estimation of mycophenolic-acid exposure from limited concentrations demonstrates that sparse TDM data can contain enough information to estimate clinically relevant exposure quantities in a well-defined application [25]. Work in neonatal vancomycin dosing likewise shows the distinction between a priori information available before drug administration and exposure estimates that can be refined once concentration information exists [26]. Clinical simulation of machine-learning-supported levothyroxine dosing additionally illustrates that the usefulness of a prediction to clinicians is a separate question from its statistical development and validation [27]. Sampling design is equally consequential: Bayesian simulations of amikacin dosing show that the timing and number of concentrations materially influence the precision with which exposure can be inferred [28]. A measured concentration is therefore not an evidentiary trump card merely because it was obtained from the patient. Its value depends on assay validity, sampling time, model identifiability, the target exposure being estimated, and whether the observation can actually distinguish between the competing explanations.

Model provenance remains important after monitoring begins. External evaluation of tacrolimus population-PK models illustrates the need to consider how relevant a candidate model is to the target population rather than judging it only by its historical use or original publication performance [29]. The same principle should apply to learned predictors. A model developed from patients who resemble the current case, under a comparable treatment objective and data-generating process, deserves different evidentiary weight from a model being applied outside its support. Reconciliation is consequently better framed as evidence weighting than as algorithm selection: each estimate carries a provisional credibility determined by population relevance, task alignment, validation status, mechanistic plausibility, temporal proximity, and patient-specific observations.

This reasoning produces a sequential but revisable decision logic. First, the predictions must be made commensurable by confirming that they concern the same dosing objective, exposure or response quantity, and decision horizon. Second, the likely source of disagreement must be identified rather than hidden by immediate averaging. Third, the information supporting each estimate should be weighted according to validation, transportability, mechanistic coherence, and relevance to the current patient. Fourth, informative TDM or response evidence should update both learned and mechanistic expectations rather than merely confirm the initially favored approach. Finally, the selected dose should remain provisional whenever decisive evidence is absent. The reconciliation rules across pre-dose prediction, observed response, therapeutic drug monitoring, and changing patient context are summarized in **Table 2**.

Table 2. Dose-reconciliation rules as patient evidence accumulates

Decision state	Information currently available	Main reconciliation question	Appropriate evidentiary operation	Dose-decision consequence	Required monitoring or revision condition
Pre-dose concordance	Learned estimate and credible PK/PD estimate point toward similar exposure or dose	Is agreement independent, or produced by shared assumptions or data?	Check target alignment, external validity and common sources of error	Concordance may support action but does not itself validate either model	Reassess after informative exposure, response, toxicity, or physiological change
Pre-dose discordance	Learned and mechanistic estimates imply materially different dosing actions	Which assumptions differ, and which are testable before treatment?	Diagnose training support, structural assumptions, population relevance and uncertainty	Select a provisional dose using the better-supported evidence; do not average automatically	Define the earliest informative observation capable of reducing uncertainty
Early post-dose observation	One or more concentrations or early response measurements become available	Does new evidence discriminate between previously plausible predictions?	Evaluate assay quality, timing, informativeness and consistency with both models	Update relative credibility and, when justified, the dose	Repeat measurement if the observation is weakly informative or discordant with both models
Evolving patient context	Physiology, organ support, interacting therapy, adherence, disease activity or clinical goals change	Have earlier model assumptions become obsolete?	Update covariates, reconsider training support and re-evaluate mechanistic structure	Re-open previously resolved decisions when the clinical state changes	Specify contextual triggers for recalculation or model re-selection
Persistent cross-model discordance	Repeated evidence does not clearly support either prediction	Is uncertainty due to insufficient data, non-identifiability, incompatible targets, or shared misspecification?	Preserve competing explanations and seek discriminating evidence	Use a cautious, reversible action when clinically feasible rather than manufacture certainty	Intensify monitoring and explicitly document the unresolved source
Post-monitoring convergence	Updated learned and mechanistic estimates become compatible with informative patient evidence	Has concordance emerged because relevant uncertainty was reduced?	Confirm that the observed agreement is robust to plausible alternative models	Continue the supported dose while retaining normal reassessment	Re-open the decision when new evidence or context invalidates the present assumptions

The resulting relationship among prediction sources is not hierarchical. Learned prediction contributes flexible pattern recognition; mechanistic PK/PD contributes structured pharmacological constraints and extrapolative reasoning; TDM contributes temporally proximal patient evidence. Their relative evidentiary value changes across the treatment course.

Proposed Dose-Reconciliation Framework

The preceding analysis suggests that reconciliation should begin before numerical disagreement itself is interpreted. Two predictions are meaningfully discordant only if they refer to the same clinical decision, exposure or response target, dosing horizon, and patient state. A learned estimate of future concentration, a mechanistic estimate of steady-state exposure, and a therapeutic target derived from a different sampling interval may appear inconsistent while answering different questions. The first operation in the proposed framework is therefore a comparability check: define what each model predicts, the assumptions under which that prediction was produced, the population and data on which it depends, and the uncertainty associated with its application to the current patient. This step prevents false reconciliation of quantities that should not have been compared.

Agreement also requires qualification. Concordant learned and mechanistic estimates may increase confidence when the approaches rely on sufficiently independent information and both remain applicable to the patient. Agreement is weaker evidence when models share development data, covariates, structural assumptions, imputed measurements, or an incorrectly specified therapeutic target. Benchmarking and validation principles developed for predictive AI and hybrid pharmacometric models therefore remain relevant even when two systems produce similar numerical answers. Concordance should be interpreted as provisional support for a decision, not as independent validation of the underlying models.

Discordance requires a different response. Rather than averaging competing estimates immediately, the framework asks which observable evidence could discriminate among the causes identified in Section 2. A prediction lying outside learned training support may deserve less weight than an independently validated mechanistic estimate, whereas a mechanistic model that poorly represents the patient's physiology may deserve less weight than a learned estimate validated in a closely matched clinical population. Patient-specific physiological information can strengthen the mechanistic side of this comparison. Physiologically based pharmacokinetic virtual-twin work in pediatric fludarabine illustrates how individual biological characteristics can be incorporated into a structured mechanistic representation and then assessed against observed exposure [30]. The importance of such information lies not in proving that PBPK is generally superior to machine learning, but in showing that pharmacological priors themselves can become more patient-specific.

The framework changes again once valid patient-linked observations become available. Prospective tacrolimus dosing provides an instructive example: an electronic-patient-file-embedded model-informed precision-dosing system used new concentration information to recalculate dosing and improved exposure-target attainment relative to physician dosing during the studied early post-transplant period [31]. That result does not establish superior clinical outcomes for all applications of automated dosing, but it demonstrates the practical value of a closed update loop in which new measurements can revise an earlier dose estimate. Under the present framework, therapeutic drug monitoring therefore functions as adjudicating evidence only to the extent that the observation is informative about the competing predictions.

This distinction produces three clinically relevant states. In provisional concordance, learned and mechanistic estimates are sufficiently compatible and adequately supported to justify action, but the decision remains open to revision as patient evidence accumulates. In discordance, the estimates imply meaningfully different actions and the cause of disagreement is sufficiently defined that discriminating evidence can be sought. In an unresolved state, available observations cannot yet establish which prediction is more credible, or both predictions appear inconsistent with emerging patient data. Practical TDM and model-informed dosing guidance for oral targeted anticancer therapy illustrates why exposure targets, sampling strategies, drug-specific evidence, and workflow must remain connected when translating monitored concentrations into dosing decisions [32]. The unresolved state therefore should not be forced into a false binary choice. Where clinically feasible, it calls for a cautious and reversible dose decision, explicit acknowledgment of residual uncertainty, and a prespecified monitoring or revision condition.

The framework consequently does not assign fixed numerical weights to AI, mechanistic PK/PD, or TDM. Their relative influence changes with the question being asked and the evidence available at that time. A learned model may dominate an initial comparison when its target population closely matches the patient and the mechanistic alternative is poorly transported. A mechanistic model may become more credible when the current physiology is interpretable but underrepresented in the learned data. Informative TDM can shift either judgment after treatment begins. If new measurements remain compatible with several explanations, uncertainty should remain visible rather than being concealed by averaging or deterministic software output.

Clinical and Methodological Implications

The framework has a direct implication for clinical implementation: a reconciliation strategy cannot be considered reliable solely because its constituent models have been published or because one produces lower retrospective error. Recommendations for integrating population-PK modeling with precision dosing emphasize the connected importance of model development, external evaluation, software implementation, workflow integration, and ongoing clinical use [33]. Under a conflict-aware approach, these stages also determine whether a disagreement can be interpreted. A prediction generated by a technically strong model may deserve little clinical weight if the target population, therapeutic objective, sampling process, or software implementation differs materially from the conditions under which that model was validated.

Prospective evaluation should therefore move beyond competitions based only on mean prediction error. Validation can be organized into progressively more demanding questions: whether each model predicts the intended quantity accurately; whether conflict between models can be recognized and correctly characterized; whether patient-linked evidence changes recommendations in the expected direction; whether the reconciliation strategy improves exposure or pharmacodynamic target attainment; and, ultimately, whether those improvements translate into patient-relevant benefit without disproportionate monitoring burden or treatment instability. The tacrolimus trial discussed above shows that prospective comparison of exposure-target attainment is feasible, but exposure improvement must remain distinct from proof of broader clinical benefit.

Clinical systems should also make disagreement auditable. When an AI recommendation and a mechanistic estimate diverge, a clinician should be able to determine the prediction target, model provenance, applicable population, major assumptions, available uncertainty information, TDM status, and reason that one estimate is being weighted more heavily. An unexplained ensemble score or automated recommendation is insufficient for a framework whose central purpose is to reason through

disagreement. This requirement also applies when estimates agree: apparent concordance should not prevent recognition of common data errors, shared assumptions, or a poorly specified therapeutic target.

Methodologically, future studies should deliberately enrich evaluation datasets for the cases in which reconciliation matters most. Average performance in uncomplicated, well-represented patients provides little evidence about the value of conflict-aware decision support. Stronger tests would identify patients in whom learned and mechanistic predictions diverge prospectively, classify the suspected source of disagreement before outcomes are known, and examine whether subsequent concentrations, pharmacodynamic observations, adverse effects, or physiological changes support one interpretation, both, or neither. Such designs would make the theoretical framework falsifiable rather than merely descriptive.

Several additional research questions follow. Investigators should determine whether model disagreement carries information about out-of-distribution patients, misspecified mechanistic priors, or changing physiology; whether predefined reconciliation rules reduce inappropriate clinician overrides; whether the value of TDM is greatest in cases with large but interpretable disagreement; and whether explicit unresolved states prevent premature confidence without producing unnecessary monitoring. Comparative studies should also test fixed AI-first, mechanism-first, simple averaging, model-selection, and conflict-aware updating strategies under identical clinical targets. No assumption should be made that the most computationally complex strategy will perform best.

The same logic has implications for governance. Updating models with local clinical data may improve relevance, but it can also introduce silent drift if changes are not evaluated against stable external criteria. Likewise, a mechanistic model can become outdated when the treated population, co-therapies, assays, or clinical objectives change. Governance should therefore focus less on whether a system is labelled “AI” or “mechanistic” and more on whether its assumptions, validation domain, changes over time, and interaction with patient observations remain visible. The core methodological requirement is not permanent agreement among models. It is the ability to recognize when disagreement is meaningful, identify what evidence could reduce it, and preserve uncertainty when available evidence cannot.

Conclusion

When a learned dose prediction conflicts with pharmacological knowledge, the discrepancy should not be resolved by assuming that either data-driven prediction or mechanistic PK/PD possesses inherent priority. Both can fail through different pathways, and both can remain locally credible despite numerical disagreement. The appropriate response is to determine whether the predictions are genuinely comparable, identify the likely source of conflict, evaluate the applicability and validation of each information source, and update the decision when informative patient-linked evidence becomes available.

Therapeutic drug monitoring is central to this process because it can provide observations from the treated patient that were unavailable at the initial dose decision. Its value, however, depends on sampling, measurement quality, target definition, and the capacity of the observation to discriminate among competing explanations. Agreement after monitoring can increase confidence; persistent disagreement may instead reveal model misspecification, population mismatch, evolving biology, or insufficiently informative data.

The proposed reconciliation framework therefore treats dose selection as a revisable evidentiary process with provisional concordant, discordant, and unresolved states. Its central claim is conditional: the safest interpretation of model conflict is not to select a permanently preferred modeling paradigm, but to make explicit which assumptions remain credible for the present patient, what evidence could change that judgment, and when uncertainty should remain visible in the clinical decision.

Acknowledgments: None

Conflict of interest: None

Financial support: None

Ethics statement: None

References

1. Huang S, Xu Q, Yang G, Ding J, Pei Q. Machine learning for prediction of drug concentrations: Application and challenges. Clin Pharmacol Ther. 2025;117(5):1236-47. doi:10.1002/cpt.3577
2. Li QY, Tang BH, Wu YE, Yao BF, Zhang W, Zheng Y, et al. Machine learning: A new approach for dose individualization. Clin Pharmacol Ther. 2024;115(4):727-44. doi:10.1002/cpt.3049
3. Minichmayr IK, Dreesen E, Centanni M, Wang Z, Hoffert Y, Friberg LE, et al. Model-informed precision dosing: State of the art and future perspectives. Adv Drug Deliv Rev. 2024;215:115421. doi:10.1016/j.addr.2024.115421
4. Darwich AS, Polasek TM, Aronson JK, Ogungbenro K, Wright DFB, Achour B, et al. Model-informed precision dosing: Background, requirements, validation, implementation, and forward trajectory of individualizing drug therapy. Annu Rev Pharmacol Toxicol. 2021;61:225-45. doi:10.1146/annurev-pharmtox-033020-113257

5. Keizer RJ, ter Heine R, Frymoyer A, Lesko LJ, Mangat R, Goswami S. Model-informed precision dosing at the bedside: Scientific challenges and opportunities. CPT Pharmacometrics Syst Pharmacol. 2018;7(12):785-7. doi:10.1002/psp4.12353
6. Dibbets AC, Koldeweij C, Osinga EP, Scheepers HCJ, de Wildt SN. Barriers and facilitators for bringing model-informed precision dosing to the patient's bedside: A systematic review. Clin Pharmacol Ther. 2025;117(3):633-45. doi:10.1002/cpt.3510
7. Ponce-Bobadilla AV, Bräm D, Farnoud A, Fröhlich H, Janssen A, Korsbo N, et al. Predictive AI in clinical pharmacology: A call to action to develop robust benchmarking practices. CPT Pharmacometrics Syst Pharmacol. 2026;15(1):e70155. doi:10.1002/psp4.70155
8. Relouw FJA, Lokers JJMT, Preijers T, Sassen SDT, van Riel NAW, Koch BCP. Beyond traditional covariates: An interpretable machine learning workflow for improved hybrid pharmacometric modeling. CPT Pharmacometrics Syst Pharmacol. 2026;15(8):e70313. doi:10.1002/psp4.70313
9. Destere A, Marquet P, Salmon Gandonnière C, Åsberg A, Loustaud-Ratti V, Carrier P, et al. A hybrid model associating population pharmacokinetics with machine learning: A case study with iohexol clearance estimation. Clin Pharmacokinet. 2022;61(8):1157-65. doi:10.1007/s40262-022-01138-x
10. Jia X, Teutonico D, Dhakal S, Psarellis YM, Abos A, Zhu H, et al. Application of machine learning and mechanistic modeling to predict intravenous pharmacokinetic profiles in humans. J Med Chem. 2025;68(7):7737-50. doi:10.1021/acs.jmedchem.5c00340
11. Jost F, Giegerich C, Grebner C, Matter H, Cordes H. Machine learning meets pharmacokinetics: A comparative analysis of predictive models for plasma concentration-time profiles. CPT Pharmacometrics Syst Pharmacol. 2026;15(5):e70226. doi:10.1002/psp4.70226
12. Teplytska O, Ernst M, Koltermann LM, Valderrama D, Trunz E, Vaisband M, et al. Machine learning methods for precision dosing in anticancer drug therapy: A scoping review. Clin Pharmacokinet. 2024;63(9):1221-37. doi:10.1007/s40262-024-01409-9
13. Agema BC, Kocher T, Öztürk AB, Giraud EL, van Erp NP, de Winter BCM, et al. Selecting the best pharmacokinetic models for a priori model-informed precision dosing with model ensembling. Clin Pharmacokinet. 2024;63(10):1449-61. doi:10.1007/s40262-024-01425-9
14. van Os W, O'Jeanson A, Troisi C, Liu C, Brooks JT, Hughes JH, et al. Machine learning-based model selection and averaging outperform single-model approaches for a priori vancomycin precision dosing. CPT Pharmacometrics Syst Pharmacol. 2025;14(10):1650-60. doi:10.1002/psp4.70084
15. Leiwolf M, Gregoire N, Magréault S, Franck B, Krekounian O, Woillard JB, et al. Model ensembling and machine learning approaches to predict the first dose of amoxicillin in intensive care. CPT Pharmacometrics Syst Pharmacol. 2026;15(7):e70296. doi:10.1002/psp4.70296
16. Le Louedec F, Morvan L, Mourey L, Maillard M, Vachoux C, Yakoubi M, et al. Model-informed precision dosing of protein kinase inhibitors: Evaluation of model-averaging and flattened priors methods. Clin Pharmacokinet. 2026;65(5):681-92. doi:10.1007/s40262-026-01628-2
17. Fochesato A, Brooks L, Bazgir O, Pierrillas PB, Jamois C, Lu J, et al. Building hybrid pharmacometric-machine learning models in oncology drug development: Current state and recommendations. CPT Pharmacometrics Syst Pharmacol. 2026;15(1):e70113. doi:10.1002/psp4.70113
18. Baran SW, Gaburro S. Hybrid mechanistic-machine learning PK/PD models with digital biomarkers: From cage to clinic. Front Pharmacol. 2026;17:1815118. doi:10.3389/fphar.2026.1815118
19. Lu J, Bender B, Jin JY, Guan Y. Deep learning prediction of patient response time course from early data via neural-pharmacokinetic/pharmacodynamic modelling. Nat Mach Intell. 2021;3(8):696-704. doi:10.1038/s42256-021-00357-4
20. Hughes JH, Tong DMH, Burns V, Daly B, Razavi P, Boelens JJ, et al. Clinical decision support for chemotherapy-induced neutropenia using a hybrid pharmacodynamic/machine learning model. CPT Pharmacometrics Syst Pharmacol. 2023;12(11):1764-76. doi:10.1002/psp4.13019
21. Wicha SG, Märtson AG, Nielsen EI, Koch BCP, Friberg LE, Alffenaar JW, et al. From therapeutic drug monitoring to model-informed precision dosing for antibiotics. Clin Pharmacol Ther. 2021;109(4):928-41. doi:10.1002/cpt.2202
22. Kluwe F, Michelet R, Mueller-Schoell A, Maier C, Klopp-Schulze L, van Dyk M, et al. Perspectives on model-informed precision dosing in the digital health era: Challenges, opportunities, and recommendations. Clin Pharmacol Ther. 2021;109(1):29-36. doi:10.1002/cpt.2049
23. Kantasiripitak W, Van Daele R, Gijsen M, Ferrante M, Spriet I, Dreesen E. Software tools for model-informed precision dosing: How well do they satisfy the needs? Front Pharmacol. 2020;11:620. doi:10.3389/fphar.2020.00620
24. Maier C, de Wiljes J, Hartung N, Kloft C, Huisinga W. A continued learning approach for model-informed precision dosing: Updating models in clinical practice. CPT Pharmacometrics Syst Pharmacol. 2022;11(2):185-98. doi:10.1002/psp4.12745
25. Woillard JB, Labriffe M, Debord J, Marquet P. Mycophenolic acid exposure prediction using machine learning. Clin Pharmacol Ther. 2021;110(2):370-9. doi:10.1002/cpt.2216
26. Tang BH, Zhang JY, Allegaert K, Hao GX, Yao BF, Leroux S, et al. Use of machine learning for dosage individualization of vancomycin in neonates. Clin Pharmacokinet. 2023;62(8):1105-16. doi:10.1007/s40262-023-01265-z

27. Janssen Daalen JM, Doesburg D, Hunik L, Kessel R, Herngreen T, Knol D, et al. Model-informed precision dosing using machine learning for levothyroxine in general practice: Development, validation and clinical simulation trial. Clin Pharmacol Ther. 2024;116(3):824-33. doi:10.1002/cpt.3293
28. Yamada T, Oda K, Nishihara M, Neo M. A simulation study on model-informed precision dosing of amikacin for achieving target area under the concentration-time curve. Br J Clin Pharmacol. 2024;90(4):1173-82. doi:10.1111/bcp.16002
29. Baroudi S, Franck B, Woillard JB, Béni NB, Comets E, Lemaitre F. Relevance-based score for MIPD model selection: Systematic review and external evaluation of popPK tacrolimus models in adult transplant recipients. CPT Pharmacometrics Syst Pharmacol. 2026;15(8):e70315. doi:10.1002/psp4.70315
30. Mseddi M, Nath C, Ben Hassine K, Mlakar V, Gloor Y, Gonzales F, et al. Physiologically based pharmacokinetic virtual twin approach for fludarabine dosing in pediatric hematopoietic stem cell transplantation. CPT Pharmacometrics Syst Pharmacol. 2026;15(3):e70218. doi:10.1002/psp4.70218
31. Kuypers DRJ, Annaert P, Jallah B, Naesens M, Teuns M, Laenen A, et al. Electronic patient file-embedded model-informed precision dosing compared with physician dosing of tacrolimus in kidney transplantation. Clin Pharmacol Ther. 2026;119(2):381-92. doi:10.1002/cpt.70090
32. Torrent-Rodríguez A, Bastida C, Carcelero San Martín E, Monge Escartín I, Puszkiel A, Blanchet B, et al. Therapeutic drug monitoring and model-informed precision dosing of oral TKIs and PARP inhibitors: A practical framework for clinical implementation. Clin Pharmacokinet. 2026;65(6):773-93. doi:10.1007/s40262-026-01660-2
33. Berezowska M, Hayden IS, Brandon AM, Zats A, Patel M, Barnett S, et al. Recommended approaches for integration of population pharmacokinetic modelling with precision dosing in clinical practice. Br J Clin Pharmacol. 2025;91(4):1064-79. doi:10.1111/bcp.16335