Pharmacophore

ISSN-2229-5402

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

INTERPRETABLE MODELS FOR PHARMACOKINETIC INTERACTION MAGNITUDE USING ENZYME, TRANSPORTER, AND DOSE FEATURES

Ahmed Benali[1]*, Karim Boudiaf[2], Samir Touati[1]

1. ***Department of Computational Drug Sciences, Faculty of Pharmacy, University of Algiers, Algiers, Algeria.***
2. ***Department of Pharmaceutical AI Systems, Faculty of Medicine, University of Tunis El Manar, Tunis, Tunisia.***

ARTICLE INFO

Received:
12 June 2025
Received in revised form:
01 October 2025
Accepted:
03 October 2025
Available online:
28 October 2025

Keywords: Explainable artificial intelligence, Pharmacokinetics, Drug–drug interactions, SHAP, CYP enzymes, Transporters

ABSTRACT

Pharmacokinetic drug interactions can significantly alter drug exposure, particularly when a perpetrator drug inhibits or induces enzymes and transporters responsible for a victim drug's clearance, and predicting the resulting fold-change in exposure requires consideration of pathway contribution, inhibitor potency, transporter involvement, and clinically relevant dose. Existing prediction tools range from simplified mechanistic equations to complex physiologically based pharmacokinetic simulations; while valuable, their outputs can be difficult for clinicians and regulators to interpret when multiple enzyme, transporter, and dose-related mechanisms act simultaneously. This article proposes an interpretable machine learning framework for predicting the AUC ratio of a victim drug when co-administered with a perpetrator drug, aiming to transparently attribute each prediction to enzyme inhibition, transporter effects, victim-drug disposition features, and perpetrator dose. A gradient-boosted tree model would be trained on curated clinical DDI evidence using features that encode CYP and transporter inhibition constants, fraction metabolized or transported, and dose-related exposure surrogates, with SHAP explanations decomposing each prediction into additive feature contributions reviewable at the level of a single interaction. Conceptually, the model would provide both a predicted interaction magnitude and a mechanistic explanation, indicating whether the prediction is primarily driven by strong CYP inhibition, transporter effects, dose-related exposure, or victim-drug pathway dependence, thereby creating a transparent audit trail. By connecting predicted exposure changes to mechanistic features, such an interpretable DDI prediction model could enhance confidence in computational pharmacokinetic risk assessment and support clinical decision-making, prescribing guidance, and regulatory review.

To Cite This Article: Benali A, Boudiaf K, Touati S. Interpretable Models for Pharmacokinetic Interaction Magnitude Using Enzyme, Transporter, and Dose Features. Pharmacophore. 2025;16(5):43-54. https://doi.org/10.51847/x9ByggPSgq

Introduction

Pharmacokinetic drug–drug interactions are clinically important because they can change victim-drug exposure through altered metabolism, transport, or both, thereby affecting efficacy and safety. Quantitative prediction of interaction magnitude has become central to drug development, labeling, and prescribing decisions, particularly when CYP enzymes, transporters, and clinically relevant perpetrator concentrations converge in the same interaction scenario [1]. Reviews of clinical DDI methodology emphasize that interpreting exposure changes requires careful attention to study design, perpetrator selection, dose, timing, and the mechanistic plausibility of the observed AUC or Cmax change [1]. Regulatory-facing pharmacokinetic assessment therefore requires not only a numerical estimate of exposure change but also a defensible explanation of why that estimate is expected.

Current approaches for DDI magnitude prediction include mechanistic static models, PBPK simulations, and increasingly data-driven models, each with a distinct balance between mechanistic detail and interpretability. PBPK models can represent tissue compartments, enzyme kinetics, transporter processes, and time-varying exposure, but their interpretation depends on extensive assumptions about physiology, drug-specific parameters, and model verification [2]. Literature on PBPK credibility

Corresponding Author: Ahmed Benali; Department of Computational Drug Sciences, Faculty of Pharmacy, University of Algiers, Algiers, Algeria. E-mail: ahmed.benali@gmail.com

also notes the need to verify model inputs and outputs, because highly detailed simulations may appear mechanistic while still being sensitive to uncertain assumptions [3]. Machine learning models could complement these tools, but their value in pharmacokinetic DDI prediction depends on making their reasoning auditable rather than merely predictive [4, 5].

Explainable machine learning offers a framework for converting complex model outputs into transparent, feature-level explanations. SHAP methods are especially relevant because they assign feature contributions to individual predictions, enabling local explanations that can be inspected by domain experts [6]. Broader XAI taxonomies emphasize that explanations should be understandable, faithful to the model, and actionable for the decision context in which the model is used [7]. In clinical pharmacology, this creates an opportunity to align data-driven prediction with established pharmacokinetic reasoning, so that a model's explanation can be compared with known enzyme and transporter mechanisms rather than accepted as an opaque output.

The central thesis of this manuscript is that an interpretable model could predict pharmacokinetic DDI magnitude from enzyme, transporter, dose, and victim-drug disposition features while generating an itemized explanation for each predicted interaction. Such a system would not replace PBPK modeling or clinical judgment; instead, it would provide a transparent surrogate or triage layer whose explanations can be challenged against mechanistic knowledge [8]. Prior work comparing machine learning, PBPK, and population pharmacokinetic approaches supports the view that model choice should reflect the decision context, available data, and need for interpretability [4]. For DDI risk assessment, the most useful model would be one that predicts exposure change while showing whether the prediction is driven by CYP inhibition, transporter inhibition, dose, fraction metabolized, or missing and uncertain evidence.

Background

Mechanisms of Pharmacokinetic Drug–Drug Interactions

Pharmacokinetic DDIs arise when a perpetrator drug changes the absorption, distribution, metabolism, or excretion of a victim drug through enzyme inhibition, enzyme induction, transporter inhibition, or combined pathway effects. Reversible inhibition, time-dependent inhibition, and induction of CYP enzymes can alter hepatic and intestinal clearance, while transporters such as P-gp, OATP, and BCRP can modify absorption, biliary transport, renal handling, and hepatic uptake [8]. The magnitude of an interaction depends not only on perpetrator potency, such as Ki or IC_{50}, but also on perpetrator exposure at the relevant site, victim-drug fraction metabolized or transported, and the availability of parallel clearance pathways [9]. Because multiple mechanisms may operate simultaneously, an interpretable model must represent both pathway-specific effects and their non-linear combinations.

Current Methods for DDI Magnitude Prediction

Mechanistic static models and PBPK simulations commonly use inputs such as inhibitor concentration, inhibition constants, fraction metabolized, fraction transported, and route-dependent exposure assumptions to estimate DDI magnitude. Static models are attractive because they are transparent and relatively simple, but they necessarily simplify temporal concentration profiles, tissue distribution, and overlapping pathways [8]. PBPK approaches provide richer mechanistic structure and are widely discussed for transporter-mediated DDIs, but their credibility depends on the quality and verification of the biological and drug-specific assumptions used in the simulation [2, 3]. This creates a gap for interpretable machine learning models that can learn empirical relationships from curated DDI evidence while still producing explanations that resemble mechanistic reasoning.

Explainable Machine Learning and SHAP

SHAP values provide a principled way to decompose a model prediction into additive feature contributions, allowing each feature to be interpreted as pushing the prediction above or below a reference value [6]. In drug development, SHAP analysis has been presented as a practical workflow for explaining supervised machine learning predictions and translating model behavior into domain-relevant evidence [10]. Applications in pharmacokinetic modeling have also used SHAP values to infer covariate relationships and support model interpretation, showing how XAI can help bridge statistical prediction and pharmacological understanding [11]. For DDI magnitude prediction, this means that a predicted AUC ratio could be accompanied by a feature-level explanation identifying the relative influence of CYP inhibition, transporter inhibition, dose, and victim-drug susceptibility.

Features That Drive DDI Magnitude

The most pharmacologically meaningful features for DDI magnitude prediction include enzyme-specific inhibition potency, transporter inhibition potency, perpetrator dose or exposure, victim-drug fraction metabolized, victim-drug fraction transported, protein binding, and blood-to-plasma partitioning. Literature on DDI assessment emphasizes that in vitro and in vivo approaches must account for the relationship between measured inhibition parameters and clinically relevant perpetrator concentrations [9]. Transporter-focused PBPK guidance similarly highlights the importance of representing transporter substrates and inhibitors in a way that reflects site-specific exposure and pathway contribution [2]. Because these variables can interact non-linearly, an interpretable tree-based model could represent threshold-like or interaction-dependent behavior while still exposing the resulting prediction logic through SHAP explanations.

Interpretable Models in Clinical Pharmacology and Regulatory Settings

Interpretable modeling is increasingly relevant in clinical pharmacology because AI-supported decisions must be understandable to pharmacologists, clinicians, and regulatory reviewers. General arguments for interpretable models in high-stakes domains caution against relying on opaque black boxes when transparent alternatives or faithful explanations are available [12]. In pharmacometrics, machine learning has been described as promising but dependent on careful validation, appropriate problem framing, and integration with established pharmacological knowledge [13]. For DDI prediction, this implies that an acceptable XAI system should document its input features, show its prediction rationale, and allow experts to decide whether the explanation is mechanistically plausible.

Model Development Overview

High-Level Prediction Pipeline

For a given victim–perpetrator pair, the proposed pipeline would construct a feature vector from in vitro potency data, pathway involvement, victim-drug disposition characteristics, and clinically relevant perpetrator dose. A gradient-boosted tree model would then estimate the expected AUC ratio, and instance-level SHAP values would be computed to explain how each feature contributed to the predicted interaction magnitude [6]. Prior machine learning work on quantitative DDI exposure prediction demonstrates that drug labels and curated interaction information can be organized into predictive features, but an XAI-oriented version must make the feature contributions available for review rather than stopping at the prediction itself [14]. The final output would therefore combine the predicted AUC ratio, an uncertainty-aware interpretation, and a mechanistic explanation suitable for pharmacological audit.

Figure 1 presents the proposed interpretable pharmacokinetic DDI prediction workflow, linking curated clinical interaction evidence to enzyme, transporter, dose, and victim-drug features, model-based AUC-ratio prediction, SHAP explanation, pharmacological audit, and regulatory-style decision support.

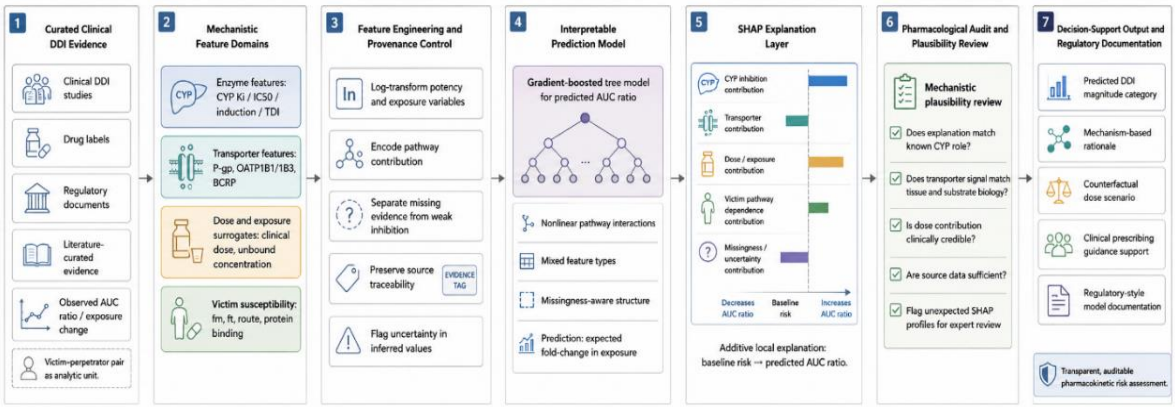

Figure 1. Interpretable Pharmacokinetic DDI Magnitude Prediction Workflow Using Enzyme, Transporter, Dose, and Victim-Drug Disposition Features

Core Input Features

The core feature set would encode enzyme features such as CYP isoform-specific Ki or IC_{50} values, transporter features such as P-gp, OATP1B1, OATP1B3, and BCRP inhibition measures, dose features such as total daily dose or unbound perpetrator concentration surrogates, and victim features such as fraction metabolized or transported. These feature classes reflect the same mechanistic concerns emphasized in reviews of in vitro and in vivo DDI assessment, where potency, exposure, and pathway contribution jointly determine interaction magnitude [9]. A model intended for clinical pharmacology should also distinguish missing evidence from evidence of no inhibition, because absent Ki data can have a different meaning from experimentally observed weak inhibition. Encoding these features explicitly would make the model's reasoning easier to compare with established mechanistic DDI expectations.

Design Principles

The model should be designed as an interpretable or explainable decision support tool rather than as an unconstrained black-box predictor. Gradient-boosted trees are attractive because they can handle mixed feature types and missingness while remaining compatible with TreeSHAP explanations, but the broader XAI literature emphasizes that explanation fidelity and human interpretability must be evaluated in the intended decision setting [7]. Clinical pharmacology applications also require robust behavior under sparse data, because many compounds lack complete transporter, enzyme, or induction measurements [13]. The design should therefore combine transparent feature definitions, missingness indicators, uncertainty communication, and explanation outputs that can be inspected by clinical pharmacologists before use in risk assessment.

Data Sources and Feature Engineering

Compilation of a DDI Magnitude Dataset

A DDI magnitude dataset would be compiled from clinical interaction studies, drug labels, literature-curated evidence, and regulatory documents, with each record representing a victim–perpetrator pair and its reported exposure change. Prior work using label information for quantitative DDI exposure prediction illustrates how regulatory and prescribing documents can be transformed into machine learning inputs, although an interpretable version would need traceable links between each feature and its evidence source [14]. Clinical methodology reviews further stress that extracted AUC ratios must be interpreted alongside study design, perpetrator dose, sampling schedule, and confidence in the observed interaction mechanism [1]. The resulting dataset should therefore preserve both the target exposure metric and the contextual evidence needed to judge whether a learned explanation is credible.

Encoding Enzyme, Transporter, and Dose Information

For each perpetrator, enzyme and transporter inhibition features would be compiled from published in vitro and clinical pharmacology sources, with Ki or IC_{50} values mapped to specific CYP isoforms and transporter systems. Dose would be encoded both as an administered clinical dose and, where possible, as an exposure-related surrogate such as unbound systemic or enterocyte-relevant concentration, because DDI risk depends on the relationship between inhibitor concentration and potency [8, 9]. Transporter-mediated interaction guidance highlights that transporter effects may differ by tissue location and substrate dependence, so feature engineering should separate uptake and efflux mechanisms rather than collapsing them into a single transporter flag [2]. This structure allows SHAP explanations to identify whether the predicted interaction is being driven by CYP inhibition, transporter inhibition, dose-related exposure, or combined mechanisms.

Table 1 defines the mechanistic feature architecture required for an interpretable pharmacokinetic DDI model, showing how enzyme, transporter, dose, victim-drug disposition, and uncertainty features contribute to both AUC-ratio prediction and SHAP-based auditability.

Table 1. Mechanistic Feature Architecture for Interpretable Pharmacokinetic DDI Magnitude Prediction

Feature domain	Representative variables	Pharmacokinetic meaning	Expected influence on predicted AUC ratio	Explanation value for SHAP audit	Key design caution
Perpetrator enzyme inhibition	CYP isoform-specific Ki, IC_{50}, reversible inhibition indicators, time-dependent inhibition flags	Captures the perpetrator's ability to inhibit metabolic clearance pathways used by the victim drug	Stronger inhibition of a major victim clearance pathway should generally push predicted AUC ratio upward	Allows reviewers to see whether the model attributes exposure increase to a plausible CYP mechanism	Ki or IC_{50} values must be linked to assay conditions, unbound concentration assumptions, and CYP isoform specificity
Perpetrator enzyme induction	Induction indicators, Emax/EC_{50} where available, chronic dosing indicators	Represents increased enzyme expression or activity that may reduce victim-drug exposure	Induction features may lower predicted AUC ratio or offset inhibition-driven increases	Helps distinguish inhibition-dominant from induction-dominant interaction profiles	Induction is time-dependent and may be underrepresented in static feature-vector models
Transporter inhibition	P-gp, BCRP, OATP1B1, OATP1B3, renal transporter inhibition measures	Captures altered absorption, efflux, hepatic uptake, biliary transport, or renal handling	Transporter inhibition may raise or sometimes contextually alter exposure depending on victim substrate dependence and tissue site	Shows whether the model is using transporter biology rather than assigning risk only to CYP features	Uptake and efflux transporters should not be collapsed into a single generic transporter flag
Perpetrator dose and exposure	Total daily dose, maximum clinical dose, unbound systemic concentration surrogate, enterocyte-relevant exposure surrogate	Connects potency to clinically relevant exposure at the likely interaction site	Higher clinically relevant exposure relative to potency should increase predicted interaction magnitude	Makes the explanation dose-sensitive and useful for counterfactual dose-adjustment reasoning	Administered dose is an imperfect proxy for local inhibitor concentration and should be interpreted cautiously
Victim-drug metabolic dependence	Fraction metabolized by specific CYPs, parallel clearance pathways, route-	Represents how vulnerable the victim drug is to inhibition or induction of a given pathway	A victim drug highly dependent on an inhibited pathway should show a larger predicted exposure increase	Allows reviewers to confirm whether the predicted interaction is aligned with	Uncertain fm values should be encoded as uncertain rather than treated as precise mechanistic truth

	dependent disposition			victim-drug disposition biology	
Victim-drug transporter dependence	Fraction transported, substrate status for uptake or efflux transporters, renal or hepatic transporter involvement	Represents susceptibility to transporter-mediated exposure changes	Transporter-dependent victim drugs may show larger predicted exposure changes when relevant transporters are inhibited	Helps separate transporter-driven interactions from metabolism-driven interactions	Substrate labels alone may be too crude unless linked to pathway contribution or tissue relevance
Physicochemical and distribution modifiers	Protein binding, blood-to-plasma ratio, route of administration, permeability-related indicators	Modifies the relationship between measured concentrations, unbound exposure, and disposition	May moderate enzyme or transporter effects rather than act as an isolated driver	Supports interpretation of why similar potency values may yield different predicted magnitudes	These variables should not obscure the primary mechanistic interpretation
Evidence completeness and uncertainty	Missingness flags, imputation indicators, data-source confidence, provenance labels	Distinguishes unavailable evidence from evidence of no effect	Missing or inferred data may widen uncertainty or reduce confidence in the predicted mechanism	Prevents explanations from overstating certainty when key inhibition, transporter, or victim features are absent	Missing evidence must not be encoded as biological absence

Victim-Drug Features

Victim-drug features would describe susceptibility to interaction, including fraction metabolized by relevant enzymes, fraction transported by relevant uptake or efflux transporters, protein binding, blood-to-plasma ratio, and route-dependent disposition characteristics. PBPK and DDI reviews emphasize that the same perpetrator can produce different exposure changes depending on how strongly the victim drug depends on the inhibited or induced pathway [3, 8]. Feature engineering should therefore represent victim-drug pathway dependence explicitly rather than treating drug identity alone as the primary predictor. When fm or ft values are uncertain, the model should encode uncertainty or missingness so that the explanation distinguishes a mechanistically supported prediction from one driven by incomplete evidence.

Interpretable Model Architecture

Choice of Algorithm

A gradient-boosted decision tree model, such as XGBoost, would be suitable for this conceptual architecture because it can represent non-linear feature interactions while remaining compatible with efficient Shapley-value explanations. Machine learning studies of pharmacokinetic DDI prediction have evaluated regression-style approaches for exposure-change prediction, suggesting that tree-based models can be considered within a broader model-comparison framework [4, 5]. The justification for using such a model in a clinical pharmacology setting would not be that it is automatically superior to mechanistic models, but that its predictions can be paired with auditable explanations that identify the dominant enzyme, transporter, and dose features. This makes the model architecture appropriate for hypothesis support, triage, and structured review rather than unexamined automation.

Feature Vector Construction and Pre-processing

Feature preprocessing would reflect pharmacokinetic reasoning by log-transforming potency and exposure-related variables where appropriate, preserving categorical indicators for enzyme and transporter involvement, and adding missingness flags for incomplete in vitro evidence. Reviews of DDI mechanisms and assessment methods emphasize that Ki, IC_{50}, inhibitor concentration, and fraction metabolized are interpreted through ratios and pathway relationships rather than as isolated linear quantities [8, 9]. If structure-based or similarity-based estimates are used to fill missing values, the imputation source should be recorded so that later explanations do not overstate the certainty of inferred data. This approach is consistent with broader machine learning guidance in pharmacometrics, where data quality, feature provenance, and interpretability are essential for responsible model use [13].

SHAP Explanation Generation

TreeSHAP would be used to compute additive feature contributions for each interaction prediction, with the base value representing the model's average prediction and the SHAP values showing how the individual feature vector shifts the prediction. The theoretical appeal of SHAP is that local explanations can be aggregated into global understanding, enabling both single-case audit trails and population-level feature importance views [6]. Practical SHAP guidance in drug development emphasizes that explanations should be interpreted in relation to the model, the data distribution, and the scientific question

rather than treated as causal proof [10]. For DDI prediction, the explanation would therefore state which features drove the predicted AUC ratio while inviting expert review of whether those drivers match known CYP, transporter, dose, and victim-drug mechanisms.

Generating And Auditing Interaction Predictions

Global Explanation – Overall Feature Importance

A global explanation layer would summarize which features most often influence predicted DDI magnitude across the curated interaction space, allowing reviewers to determine whether the model's behavior is pharmacologically credible. If CYP3A4 inhibition, perpetrator exposure, P-gp inhibition, OATP involvement, and victim-drug pathway dependence emerge as influential features, that pattern would be expected to align with established DDI mechanisms and transporter-mediated PBPK principles [2]. Similar explainability goals appear in interpretable DDI modeling studies that use biological or molecular features to clarify why particular drug pairs are predicted to interact [15]. Global SHAP summaries would therefore serve as a model-level plausibility check rather than as a substitute for local, interaction-specific review.

Local Explanation – Per-Interaction Audit Trail

For an individual victim–perpetrator pair, the local explanation would show how the feature vector moves the prediction from the model baseline toward a higher or lower expected AUC ratio. A SHAP waterfall or force-style explanation could state that the prediction is mainly supported by strong inhibition of the victim drug's major CYP pathway, with additional support from transporter inhibition and clinically relevant perpetrator dose, while weak or absent features contribute little to the prediction [6]. Explainable graph-based and gene-expression-based DDI studies illustrate the broader value of moving beyond an interaction flag toward explanations that identify the biological or pharmacological basis for a prediction [15, 16]. In the proposed pharmacokinetic model, the audit trail would allow a clinical pharmacologist to verify whether the explanation is consistent with known enzyme and transporter roles.

Detecting Unexpected Predictions

Unexpected predictions would be flagged when the model assigns a large predicted exposure change to a mechanism that appears inconsistent with established pharmacology, such as a minor transporter feature dominating a prediction for a victim drug usually cleared by metabolism. Reviews of machine learning approaches for DDI prediction emphasize that data-driven systems can inherit biases, sparsity, or annotation errors from source datasets, making expert review essential when predictions appear mechanistically surprising [17]. Multimodal and deep learning DDI frameworks show that complex models can integrate diverse drug features, but such flexibility can also make transparent review more important when the model detects patterns not obvious from classical mechanisms [18]. In this framework, unexpected SHAP profiles would trigger review of the input data, feature encoding, and source evidence before the prediction is used for decision support.

Counterfactual Analysis for Dose Adjustment

Counterfactual analysis would allow reviewers to modify one or more inputs, such as perpetrator dose or exposure surrogate, and observe how the predicted AUC ratio and its explanation would be expected to change. This is particularly relevant because pharmacokinetic interaction magnitude depends on the relationship between inhibitor concentration, potency, and pathway contribution, not simply on whether a drug is labeled as an inhibitor [8]. Machine learning models developed for DDI prediction increasingly incorporate molecular, network, and pharmacological features, but dose-sensitive pharmacokinetic interpretation requires explicit representation of exposure-related inputs [19, 20]. A counterfactual explanation could therefore support dose-adjustment reasoning by showing whether reducing the perpetrator exposure would be expected to reduce the contribution from a specific enzyme or transporter mechanism.

Explainability for Clinical and Regulatory Review

Explanation Interface for Clinical Pharmacologists

An explanation interface for clinical pharmacologists would display the predicted exposure-change category, the SHAP decomposition, the relevant Ki or IC_{50} values, victim-drug pathway dependence, and the evidence source for each input. Such an interface would be consistent with the broader expectation that machine learning in pharmacometrics should be interpretable, scientifically grounded, and integrated into expert workflows rather than treated as autonomous decision-making [13]. Prior DDI prediction models using structural similarity, interaction networks, and pharmacokinetic or pharmacodynamic knowledge show that prediction outputs become more useful when linked back to interpretable drug properties [21]. For clinical review, the interface should make it easy to ask whether the model's stated rationale matches the known pharmacology of the pair.

Regulatory-Style Model Documentation

Regulatory-style documentation would describe the model objective, intended use, feature definitions, data provenance, preprocessing choices, explanation method, validation plan, and limitations in a format suitable for model-informed drug development review. Regulatory alignment should be treated as a core design requirement rather than a final reporting step. The ICH M12 guideline now provides a harmonized framework for designing, conducting, and interpreting enzyme- and

transporter-mediated DDI studies, while FDA and EMA guidance documents define practical expectations for clinical and in vitro DDI assessment, including CYP inhibition or induction, transporter liability, study interpretation, and labeling relevance [22–25]. Earlier regulatory-science work also shows that DDI evaluation has evolved from isolated CYP-focused assessment toward integrated consideration of enzymes, transporters, exposure, and decision-ready labeling language [26]. For an interpretable machine-learning model, this means that feature definitions, evidence provenance, SHAP explanations, and uncertainty flags should be organized in a way that can be compared with regulatory DDI logic, PBPK review practices, and transporter-science recommendations [27, 28].

PBPK transporter DDI recommendations and PBPK credibility discussions both emphasize that model assumptions, parameter sources, and verification steps must be transparent when computational outputs inform drug-development decisions [2, 3]. Explainable AI principles add that the reasoning process should be understandable to the human decision-maker, especially in high-stakes settings where opaque prediction is insufficient [7, 12]. The documentation would therefore present both predictive logic and explanation logic, enabling reviewers to examine how enzyme, transporter, and dose features shape model output.

Auditing Model Predictions against Known Interactions

Auditing would involve applying the model to well-characterized pharmacokinetic interactions and comparing the generated explanations with accepted mechanistic understanding. If a known CYP-mediated interaction is explained mainly by CYP inhibition and victim-drug fraction metabolized, the model's rationale would be considered mechanistically plausible, whereas a transporter-dominated explanation would require further review [1]. Reviews of DDI prediction based on deep learning and knowledge graphs show that modern systems can identify complex interaction patterns, but they also reinforce the need to connect predictions to interpretable mechanisms when the output may affect therapeutic decisions [29]. The purpose of this audit would be to evaluate explanation plausibility, not merely whether the predicted magnitude appears numerically close to an observed value.

Feedback Loop for Model Refinement

A feedback loop would allow pharmacologists to flag explanations that appear implausible, incomplete, or inconsistent with source evidence, and those flags would guide feature revision, data curation, or additional mechanistic review. Critical reviews of machine learning-based DDI prediction highlight recurring challenges such as heterogeneous data sources, sparse annotations, and limited external validation, all of which can affect both predictions and explanations [30, 31]. Mechanism-focused DDI resources and models show the value of linking predicted interactions to clarified mechanisms, because mechanistic annotation can improve the interpretability of downstream risk assessment [32]. In this framework, expert feedback would not simply correct isolated outputs but would improve the alignment between feature engineering, model behavior, and pharmacological reasoning. **Figure 2** illustrates how pharmacologist feedback can convert questionable DDI explanations into targeted feature revision, data curation, mechanistic review, and improved alignment between model behavior and pharmacological reasoning.

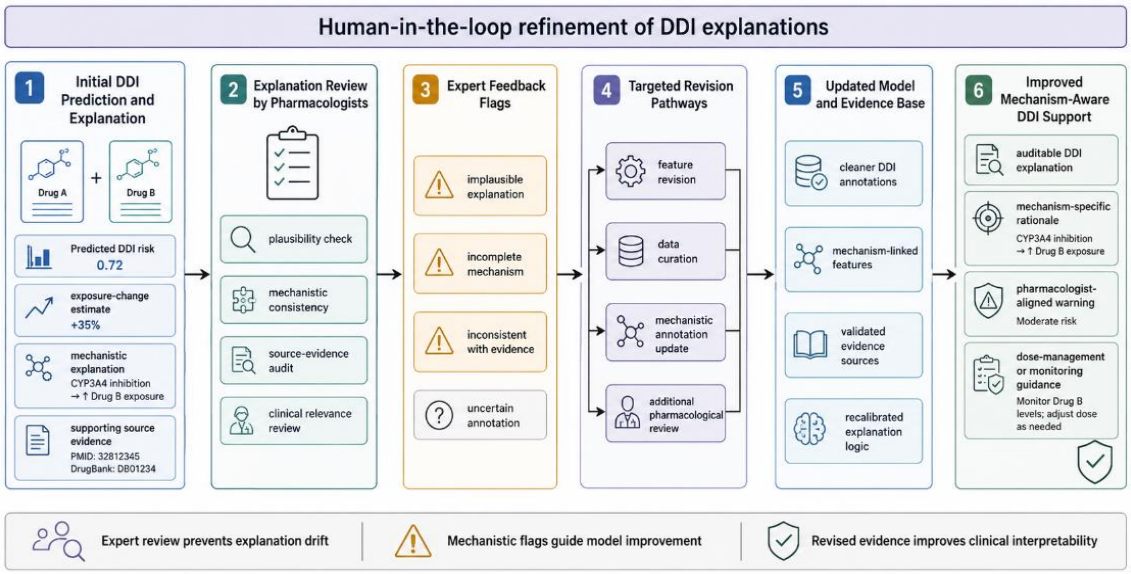

Figure 2. Expert Feedback Loop for Mechanism-Aware DDI Explanation Refinement

Integration Into DDI Risk Assessment Workflow

Use in Early-Stage Perpetrator Screening

In early discovery and development, the interpretable model could screen candidate perpetrators at anticipated clinical exposure ranges and identify which enzyme or transporter liabilities are most responsible for predicted interaction risk. This use case would complement in vitro and PBPK workflows by prioritizing mechanisms that require additional de-risking, such as potent CYP inhibition, clinically relevant transporter inhibition, or uncertainty in victim-drug pathway contribution [9]. Reviews of machine learning in pharmacometrics suggest that such models are most useful when positioned as decision-support tools that help prioritize experiments and interpret evidence rather than replace mechanistic investigation [13]. The explanation layer would make the screening output actionable by showing whether the risk is driven by potency, dose, pathway dependence, or missing information.

Augmenting Clinical Decision Support for Polypharmacy

In clinical decision support, an interpretable pharmacokinetic DDI model could evaluate a patient's regimen and return not only an interaction warning but also a transparent rationale for the expected exposure change. Polypharmacy-oriented graph models demonstrate the importance of identifying drug combinations that may cause clinically relevant effects, but pharmacokinetic prescribing guidance requires explanations that distinguish metabolic inhibition, transporter inhibition, and exposure-dependent mechanisms [33]. Deep learning approaches have improved broad DDI prediction, including drug–drug and drug–food interaction identification, yet clinical uptake would be strengthened by explanations that pharmacists and physicians can audit [20]. The proposed model would therefore translate a computational warning into a mechanism-based statement that can support medication review and dose-management decisions. **Table 2** summarizes how an interpretable pharmacokinetic DDI model can convert polypharmacy interaction prediction into mechanism-based clinical decision support.

Table 2. Clinical Use Pathway for Interpretable Pharmacokinetic DDI Alerts in Polypharmacy Review

CDS output layer	What the model should report	Clinical value for medication review
Interaction signal	Flag whether the drug pair or multi-drug combination is likely to alter exposure.	Helps prioritize clinically relevant interactions rather than generating nonspecific alerts.
Mechanistic explanation	Identify the likely driver, such as metabolic inhibition, transporter inhibition, induction, or exposure-dependent toxicity.	Allows pharmacists and physicians to judge whether the alert is pharmacologically plausible.
Exposure direction and magnitude	Indicate whether drug exposure is expected to increase or decrease, with approximate risk level.	Supports dose adjustment, substitution, monitoring, or temporary withholding decisions.
Patient-regimen context	Link the alert to the patient's active medications, dose intensity, comorbid risk factors, and overlapping toxicities.	Reduces alert fatigue by showing why the warning matters for this specific patient.
Actionable recommendation	Suggest review options such as dose reduction, therapeutic drug monitoring, alternative therapy, or closer follow-up.	Converts model output into a practical decision-support step without replacing clinician judgment.

Evaluation Strategy

Predictive Performance Metrics

Evaluation would compare predicted and observed exposure-change patterns using pre-specified pharmacokinetic performance metrics, but the manuscript would treat these as planned assessment tools rather than report any experimental results. Prior regression-oriented work in pharmacokinetic DDI prediction provides a basis for evaluating whether machine learning can support quantitative exposure prediction, while comparisons with PBPK and population pharmacokinetic methods emphasize that evaluation should be matched to the intended use case [4, 5]. Benchmarking against mechanistic static models and conventional machine learning baselines would help determine whether the interpretable model offers decision-relevant value beyond existing approaches. The key requirement is that any predictive assessment be accompanied by explanation assessment, because a numerically plausible prediction may still be unsuitable if its rationale is pharmacologically implausible.

Explanation Quality and Mechanistic Plausibility

Explanation quality would be evaluated by asking clinical pharmacologists to judge whether the top contributing features for selected predictions are consistent with known mechanisms, input evidence, and expected pathway dependence. Practical SHAP guidance emphasizes that explanations should be interpreted with scientific context, because SHAP values describe model behavior rather than proving biological causation [10]. Interpretable DDI models based on drug-induced gene expression, graph reasoning, and clarified mechanism annotations show that explanation plausibility can be framed as the consistency between model-identified drivers and domain knowledge [15, 16, 32]. This evaluation would therefore examine whether the model's explanations are useful, faithful, and mechanistically reviewable rather than simply visually appealing.

Table 3 provides a review-ready evaluation framework for interpretable pharmacokinetic DDI prediction, integrating numerical performance, SHAP explanation quality, mechanistic plausibility, uncertainty handling, counterfactual dose sensitivity, and regulatory documentation readiness.

Table 3. Explanation and Evaluation Framework for Review-Ready Pharmacokinetic DDI Prediction

Review dimension	What should be evaluated	Practical evaluation approach	What would count as a strong result	What would trigger expert review	Decision-support implication
Numerical prediction accuracy	Agreement between predicted and observed AUC ratio or exposure-change category	Compare predicted versus observed AUC ratios using regression and category-based performance metrics	Predicted magnitude is close enough to support triage, labeling review, or further pharmacokinetic assessment	Large prediction error for well-characterized interactions	Determines whether the model is quantitatively useful for DDI risk screening
Mechanistic plausibility of explanation	Whether top SHAP contributors match accepted pharmacokinetic mechanisms	Clinical pharmacologists review local SHAP profiles for known interactions	CYP-driven interactions are explained by CYP inhibition and victim fm; transporter-mediated interactions show relevant transporter features	Dominant feature contribution contradicts known clearance or transporter biology	Determines whether the output can be trusted as more than a black-box prediction
Local explanation completeness	Whether the explanation includes the main enzyme, transporter, dose, and victim-drug features relevant to a specific pair	Inspect per-interaction SHAP waterfall or contribution table	Explanation identifies the principal mechanism and clinically meaningful secondary contributors	Important known mechanism is absent or assigned negligible contribution	Supports single-case audit trails for clinicians, pharmacologists, and reviewers
Global model behavior	Whether population-level feature importance aligns with pharmacological expectations	Review global SHAP summaries across the curated interaction dataset	CYP3A4 inhibition, perpetrator exposure, transporter involvement, and victim pathway dependence emerge as influential where expected	Non-mechanistic or provenance-related variables dominate global importance	Supports model-level credibility before deployment or regulatory-facing use
Uncertainty and missingness handling	Whether incomplete Ki, IC_{50}, transporter, induction, or fm evidence is visible in the output	Compare predictions with and without missingness indicators or imputed features	The interface clearly distinguishes measured evidence from inferred or absent evidence	Prediction appears confident despite sparse or inferred mechanistic inputs	Prevents overconfident interpretation of poorly supported DDI predictions
Counterfactual dose sensitivity	Whether changing dose or exposure surrogate changes prediction and explanation plausibly	Modify perpetrator dose or exposure inputs and observe predicted AUC ratio and SHAP changes	Lower exposure reduces dose-driven contribution when pharmacologically expected	Prediction remains unchanged despite major clinically relevant dose changes	Supports dose-adjustment reasoning and scenario testing
Temporal and external generalizability	Whether predictions remain valid for newer drugs, new labels, and emerging interaction evidence	Use time-split validation with older evidence for training and newer evidence for external evaluation	Model maintains prediction quality and plausible explanations on more recent DDI cases	Performance drops for new transporter mechanisms, newer drug classes, or sparse evidence	Determines readiness for ongoing DDI surveillance and label-update contexts
Regulatory documentation readiness	Whether the model can be reviewed as a transparent computational tool	Document intended use, feature definitions, data provenance, preprocessing, model choice, SHAP method, validation, and limitations	Review package allows independent inspection of inputs, assumptions, prediction logic, and explanation logic	Missing provenance, unclear feature definitions, or undocumented preprocessing	Determines whether the model can support regulatory-style review rather than informal exploration
Human review integration	Whether expert users can challenge, flag, or contextualize model explanations	Pharmacologists review unexpected profiles and record explanation-quality judgments	Expert review identifies actionable data gaps, mechanistic inconsistencies, or validation needs	Users treat model output as autonomous prescribing guidance without audit	Preserves the model's role as decision support rather than replacement for clinical

	pharmacology judgment

Prospective Evaluation

Prospective evaluation would use a time-split strategy in which older interaction evidence supports model development and more recent clinical DDI reports are reserved for external review. Temporal validation is important because DDI evidence evolves as new drugs, transporter findings, and regulatory label updates appear, and reviews of DDI prediction methods warn that models trained on static datasets may not generalize to emerging interaction mechanisms [17, 30]. Machine learning frameworks for DDI prediction, including multimodal and feature-fusion approaches, further indicate that external validation is needed when models integrate heterogeneous biological, chemical, and clinical information [18, 19]. A prospective strategy would assess whether both predictions and explanations remain plausible as the evidence base changes.

Figure 3 shows how a prospective time-split validation design can test whether DDI predictions and mechanistic explanations remain reliable when evaluated on newer interaction evidence that was not available during model development.

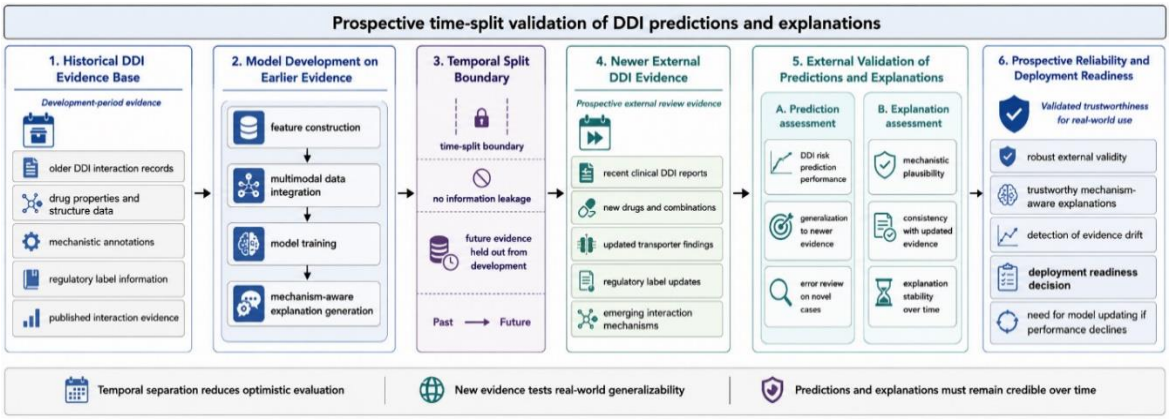

Figure 3. Prospective Time-Split External Validation Framework for DDI Prediction and Explanation Stability.

Limitations

Dependence on In Vitro Data Quality

The proposed model would depend strongly on the quality, consistency, and relevance of in vitro Ki, IC_{50}, induction, and transporter measurements. In vitro and in vivo DDI assessment reviews note that assay conditions, protein binding assumptions, and concentration selection can substantially affect how inhibition parameters are interpreted for clinical risk assessment [8, 9]. PBPK credibility work similarly emphasizes that uncertain or poorly verified inputs can propagate through a model and undermine confidence in outputs [3]. The explanation interface should therefore display input uncertainty and data provenance so that users can recognize when a prediction is supported by limited or variable evidence.

Limited Handling of Induction and Time-Dependent Inhibition

Although the model could include induction parameters and indicators for time-dependent inhibition, delayed enzyme turnover, chronic dosing, and mechanism-based inhibition may not be fully represented without richer temporal features. Clinical DDI methodology stresses that interaction magnitude can depend on dosing schedule, duration, sampling time, and whether the mechanism involves reversible inhibition, induction, or time-dependent loss of enzyme activity [1]. PBPK models remain better suited than simple feature-vector models for representing time-varying concentrations and biological turnover when detailed parameterization is available [2]. The interpretable machine learning model should therefore be positioned as a transparent decision-support and screening framework, not as a replacement for mechanistic simulation when temporal kinetics dominate the interaction.

Conclusion

An interpretable model for pharmacokinetic interaction magnitude would combine quantitative prediction with mechanism-level transparency. By using enzyme, transporter, dose, and victim-drug disposition features, the model could estimate the expected direction and magnitude of exposure change while showing which inputs drive the prediction.

The main strength of this approach is auditability. A clinical pharmacologist, pharmacist, or regulatory reviewer could inspect whether a prediction is driven by plausible CYP inhibition, transporter effects, perpetrator exposure, or victim-drug pathway dependence, rather than accepting a black-box warning.

Important challenges remain. The model would be sensitive to variable in vitro data, incomplete transporter evidence, sparse clinical interaction studies, and mechanisms such as induction or time-dependent inhibition that may require more explicit temporal representation.

Future progress will require collaboration among clinical pharmacologists, pharmacometricians, informaticians, regulators, and data curators. A gold-standard, transparent DDI dataset and review-ready explanation interfaces would help make interpretable AI a practical component of prescribing guidance and regulatory DDI assessment.

Acknowledgments: None

Conflict of interest: None

Financial support: None

Ethics statement: None

References

1. Tornio A, Filppula AM, Niemi M, Backman JT. Clinical studies on drug–drug interactions involving metabolism and transport: methodology, pitfalls, and interpretation. Clin Pharmacol Ther. 2019;105(6):1345-61.
2. Taskar KS, Pilla Reddy V, Burt H, Posada MM, Varma M, Zheng M, et al. Physiologically-based pharmacokinetic models for evaluating membrane transporter mediated drug–drug interactions: current capabilities, case studies, future opportunities, and recommendations. Clin Pharmacol Ther. 2020;107(5):1082-115.
3. Rodrigues D, Gibson CR, Isoherranen N. Drug interaction PBPK modeling: Review of the literature exposes the need for increased verification of model inputs and outputs as part of credibility assessment. Clin Transl Sci. 2025;18(7):e70299.
4. Gill J, Moullet M, Martinsson A, Miljković F, Williamson B, Arends RH, et al. Comparing the applications of machine learning, PBPK, and population pharmacokinetic models in pharmacokinetic drug–drug interaction prediction. CPT Pharmacometrics Syst Pharmacol. 2022;11(12):1560-8.
5. Gill J, Moullet M, Martinsson A, Miljković F, Williamson B, Arends RH, et al. Evaluating the performance of machine-learning regression models for pharmacokinetic drug–drug interactions. CPT Pharmacometrics Syst Pharmacol. 2023;12(1):122-34.
6. Lundberg SM, Erion G, Chen H, DeGrave A, Prutkin JM, Nair B, et al. From local explanations to global understanding with explainable AI for trees. Nat Mach Intell. 2020;2(1):56-67.
7. Barredo AA, Del Ser J, Gil-Lopez S, Díaz-Rodríguez N, Bennetot A, Chatila R, et al. Explainable explainable artificial intelligence (XAI): Concepts, taxonomies, opportunities and challenges toward responsible AI. Inf Fusion. 2020;58:82-115.
8. Peng Y, Cheng Z, Xie F. Evaluation of pharmacokinetic drug–drug interactions: a review of the mechanisms, in vitro and in silico approaches. Metabolites. 2021;11(2):75.
9. Lu C, Di L. In vitro and in vivo methods to assess pharmacokinetic drug–drug interactions in drug discovery and development. Biopharm Drug Dispos. 2020;41(1-2):3-1.
10. Ponce-Bobadilla AV, Schmitt V, Maier CS, Mensing S, Stodtmann S. Practical guide to SHAP analysis: Explaining supervised machine learning model predictions in drug development. Clin Transl Sci. 2024;17(11):e70056.
11. Janssen A, Hoogendoorn M, Cnossen MH, Mathôt RA, OPTI-CLOT Study Group, et al. Application of SHAP values for inferring the optimal functional form of covariates in pharmacokinetic modeling. CPT Pharmacometrics Syst Pharmacol. 2022;11(8):1100-10.
12. Rudin C. Stop explaining black box machine learning models for high stakes decisions and use interpretable models instead. Nat Mach Intell. 2019;1(5):206-15.
13. McComb M, Bies R, Ramanathan M. Machine learning in pharmacometrics: Opportunities and challenges. Br J Clin Pharmacol. 2022;88(4):1482-99.
14. Jang HY, Song J, Kim JH, Lee H, Kim IW, Moon B, et al. Machine learning-based quantitative prediction of drug exposure in drug-drug interactions using drug label information. NPJ Digit Med. 2022;5(1):88.
15. Kim E, Nam H. DeSIDE-DDI: interpretable prediction of drug-drug interactions using drug-induced gene expressions. J Cheminform. 2022;14(1):9.
16. Zhong Y, Chen X, Zhao Y, Chen X, Gao T, Weng Z. Graph-augmented convolutional networks on drug-drug interactions prediction. arXiv. 2019;arXiv:1912.03702.
17. Han K, Cao P, Wang Y, Xie F, Ma J, Yu M, et al. A review of approaches for predicting drug–drug interactions based on machine learning. Front Pharmacol. 2022;12:814858.
18. Deng Y, Xu X, Qiu Y, Xia J, Zhang W, Liu S. A multimodal deep learning framework for predicting drug–drug interaction events. Bioinformatics. 2020;36(15):4316-22.
19. Chen Y, Ma T, Yang X, Wang J, Song B, Zeng X. MUFFIN: multi-scale feature fusion for drug–drug interaction prediction. Bioinformatics. 2021;37(17):2651-8.
20. Ryu JY, Kim HU, Lee SY. Deep learning improves prediction of drug–drug and drug–food interactions. Proc Natl Acad Sci U S A. 2018;115(18):E4304-11.

21. Takeda T, Hao M, Cheng T, Bryant SH, Wang Y. Predicting drug–drug interactions through drug structural similarities and interaction networks incorporating pharmacokinetics and pharmacodynamics knowledge. J Cheminform. 2017;9(1):16.
22. International Council for Harmonisation. ICH M12 drug interaction studies: Final guideline. Geneva: ICH; 2024.
23. U.S. Food and Drug Administration. Clinical drug interaction studies—Cytochrome P450 enzyme- and transporter-mediated drug interactions: Guidance for industry. Silver Spring (MD): FDA; 2020.
24. U.S. Food and Drug Administration. In vitro drug interaction studies—Cytochrome P450 enzyme- and transporter-mediated drug interactions: Guidance for industry. Silver Spring (MD): FDA; 2020.
25. European Medicines Agency. Guideline on the investigation of drug interactions. CPMP/EWP/560/95/Rev. 1 Corr. 2. London: EMA; 2012.
26. Huang SM, Strong JM, Zhang L, Reynolds KS, Nallani S, Temple R, et al. New era in drug interaction evaluation: US Food and Drug Administration update on CYP enzymes, transporters, and the guidance process. J Clin Pharmacol. 2008;48(6):662-70.
27. Zhao P, Zhang L, Grillo JA, Liu Q, Bullock JM, Moon YJ, et al. Applications of physiologically based pharmacokinetic modeling and simulation during regulatory review. Clin Pharmacol Ther. 2011;89(2):259-67.
28. Giacomini KM, Huang SM, Tweedie DJ, Benet LZ, Brouwer KLR, Chu X, et al. Membrane transporters in drug development. Nat Rev Drug Discov. 2010;9(3):215-36.
29. Luo H, Yin W, Wang J, Zhang G, Liang W, Luo J, et al. Drug-drug interactions prediction based on deep learning and knowledge graph: A review. iScience. 2024;27(3).
30. Wang NN, Zhu B, Li XL, Liu S, Shi JY, Cao DS. Comprehensive review of drug–drug interaction prediction based on machine learning: current status, challenges, and opportunities. J Chem Inf Model. 2023;64(1):96-109.
31. Gheorghita FI, Bocanet VI, Iantovics LB. Machine learning-based drug-drug interaction prediction: a critical review of models, limitations, and data challenges. Front Pharmacol. 2025;16:1632775.
32. Hu W, Zhang W, Zhou Y, Luo Y, Sun X, Xu H, et al. MECDDI: clarified drug–drug interaction mechanism facilitating rational drug use and potential drug–drug interaction prediction. J Chem Inf Model. 2023;63(5):1626-36.
33. Zitnik M, Agrawal M, Leskovec J. Modeling polypharmacy side effects with graph convolutional networks. Bioinformatics. 2018;34(13):i457-66.