Title Page

Real-world application of PBPK in drug discovery

Laura G A Santos¹, Swati Jaiswal¹, Kuan-Fu Chen, Hannah M Jones, Ian E Templeton²


¹ Co-First authors

² Corresponding author
Real-world application of PBPK in drug discovery

Laura G A Santos¹, Swati Jaiswal¹, Kuan-Fu Chen, Hannah M Jones, Ian E Templeton²


¹ Co-First authors

² Corresponding author: ian.templeton@certara.com; 100 Overlook Center, Suite 101, Princeton, NJ 08540

Authorship Contributions

Participated in research design: Santos, Jaiswal, Chen, and Jones

Conducted experiments: Santos and Jaiswal

Performed data analysis: Santos, Jaiswal, Chen, and Jones

Wrote or contributed to the writing of the manuscript: Santos, Jaiswal, Chen, Jones and Templeton

The total number of manuscript pages: 37
Total number of Figures: 7
Total number of Tables: 1
Total word count of the Abstract: 238
Total word count of the Introduction: 274
Total word count of the Discussion (Conclusions): 158

This work received no external funding.
Non-standard abbreviations

ADAM: Advanced Dissolution, Absorption and Metabolism
AUC: Area Under the Plasma Concentration-Time Curve
BCS: Biopharmaceutical Classification System
BDDCS: Biopharmaceutical Drug Disposition and Classification System
B:P: Blood: Plasma Ratio
CL: Clearance
CLint: Intrinsic Clearance
Cmax: Maximum Concentration
Cmin: Minimum Concentration
DDI: Drug Drug Interaction
EC90: 90% Maximal Effective Concentration
F: Oral Bioavailability
Fa: Fraction of Oral Dose Absorbed
FIH: First in Human
fmCYP: Fraction Metabolized by Specific CYP Isoform
Fu: Unbound Fraction in Plasma
GMR: Geometric Mean Ratio
HH: Human Hepatocyte
HLM: Human Liver Microsome
ISEF: Inter System Extrapolation Factor
IV: Intravenous
IVIVC: In Vitro to In Vivo Correlation
IVIVE: In Vitro to In Vivo Extrapolation
Ka: First Order Absorption Rate
Ki: Concentration Required to Produce Half Maximum Inhibition
Kp: Tissue: Plasma Partition Coefficient
PBPK: Physiologically Based Pharmacokinetic
PD: Pharmacodynamic
PK: Pharmacokinetic
pKa: Acid Dissociation Constant
PO: Per Oral
Peff,man: Effective Permeability Predicted in Human
Ptrans,0: Intrinsic Transcellular Permeability
QD: Daily Dosing (every 24 h)
SIVA: Simcyp In Vitro Analysis
Vss: Distribution Volume at Steady State
Abstract

The utility of PBPK models in support of drug development has been well documented. During the discovery stage, PBPK has increasingly been applied for early risk assessment, prediction of human dose, toxicokinetic dose projection and early formulation assessment. Previous review articles have proposed model building and application strategies for PBPK-based first in human predictions with comprehensive descriptions of the individual components of PBPK models. This includes the generation of decision trees, based on comprehensive literature reviews, to guide the application of PBPK in the discovery setting. The goal of this mini review is to provide additional guidance on the real-world application of PBPK, in support of the discovery stage of drug development. In this mini review, our goal is to provide guidance on the typical steps involved in the development and application of a PBPK model during drug discovery to assist in decision making. We have illustrated our recommended approach through description of case examples, where PBPK has been successfully applied to aid in human PK projection, candidate selection and prediction of drug interaction liability for parent and metabolite. Through these case studies, we have highlighted fundamental issues, including pre-verification in preclinical species, the application of empirical scalars in the prediction of in vivo clearance from in vitro systems, in silico prediction of permeability and the exploration of aqueous and biorelevant solubility data to predict dissolution. In addition, current knowledge gaps have been highlighted and future directions proposed.

Significance Statement

Through description of three case studies, we have highlighted the fundamental principles of PBPK application during drug discovery. These include pre-verification of the model in preclinical species, application of empirical scalars where necessary in the prediction of clearance, in silico prediction of permeability, and the exploration of aqueous and biorelevant solubility data to predict dissolution. In addition, current knowledge gaps have been highlighted and future directions proposed.
Introduction

The goal of this mini review is to provide additional guidance on the real-world application of PBPK, in support of the discovery stage of drug development (Figure 1). During the discovery stage, the real-world evidence available is relatively consistent across development programs. Data at this stage generally includes a combination of in vitro and preclinical data. In vitro data includes information on drug metabolism, plasma protein binding, intestinal permeability, and preliminary drug interaction potential assessments. Preclinical data includes clearance and distribution volume after IV dosing, and rate and extent of absorption after oral administration. In most cases, formulations data, such as dissolution rate and pH-dependent solubility data are available at this time as well. These in vitro and preclinical data are crucial to the real-world application of PBPK in support of drug discovery.

Previous publications have focused on a more comprehensive description of the individual components of PBPK models. In this mini review, our goal is to provide guidance on the typical steps involved in developing and applying a PBPK model during drug discovery to assist in decision making. We have illustrated our recommended approach through the description of case examples, where PBPK has been successfully applied to aid in human PK projection, candidate selection and prediction of drug interaction liability for parent and metabolite. Through these case studies, we have highlighted fundamental issues, including pre-verification in preclinical species, the application of empirical scalars in the prediction of in vivo clearance from in vitro systems, in silico prediction of permeability and the exploration of aqueous and biorelevant solubility data to predict dissolution. In addition, current knowledge gaps have been highlighted and future directions proposed.
Brief Historical Perspective

Physiologically based pharmacokinetic (PBPK) models are complex compartmental models which describe tissues, or tissue subsections, which are linked together by blood flow. Each compartment is defined by system parameters specific to the species or human population of interest. These include blood flow, volume, transit time, pH, and abundance of specific drug metabolizing enzymes or transporters. The PBPK model user integrates drug-specific parameters, such as physicochemical data (e.g., pKa, Log P), unbound fraction in plasma (fu), blood-to-plasma concentration ratio (B:P), solubility, and permeability as well as intrinsic clearance from microsomes, hepatocytes and recombinant systems, to predict the absorption, distribution, and clearance of the drug in the body. Drug-specific parameters, combined with the system parameters within the PBPK model, allow prediction of PK profiles in human populations and preclinical species. In addition, by incorporating inhibition or induction parameters into the model, the user can explore how drugs may interact with each other.

The utility of PBPK models in support of drug development has been well documented. Over the past 15 years, PBPK models have been applied in both the discovery and development stages. Generally, when applied during the development stage, the application of PBPK has been focused on the prediction of drug-drug interactions, changes in pharmacokinetics across different subject populations, or the impact of changes in drug formulation on drug absorption. During the discovery stage, PBPK has increasingly been applied for early risk assessment, prediction of human dose, toxicokinetic dose projection and early formulation assessment. In a thorough review article (Miller et al., 2019) the authors proposed an enhancement to the previously proposed model building and application strategies for PBPK-based first in human predictions (Jones et al., 2006). This enhancement included the generation of decision trees to guide the application of PBPK in the discovery setting, which were based on a comprehensive review of recent literature, and the best practices of experienced PBPK users. Similarly, Mao et al. (2023) have more recently shared a thorough retrospective analysis of the benefits of PBPK support for internal compounds. The objective of the
current tutorial is to provide a practical guide for the real-world application of PBPK
during drug discovery.

Although PBPK modelling is most effectively performed over different stage gates
(discovery, early and late development) in a learn and confirm cycle, there is potential
for impact on decision making at the earliest stages of the drug development process.
Beginning in discovery, PBPK supports lead molecule identification and optimization.
PBPK can guide decisions regarding optimal laboratory objectives for a chemical series.
High throughput screening for rapid batch processing of results to triage the best drug
candidates is also supported. As development progresses toward candidate selection,
PBPK delivers early PK, pharmacodynamic (PD) and most importantly, a mechanistic
first in human (FIH) PK prediction. Additional applications include formulation
development and prediction of exposure in preclinical toxicology studies.

In general, PBPK application during discovery begins with performing a PK prediction in
preclinical species. This allows the user to build confidence that the assumptions of the
model are valid, before moving forward to the human prediction. Typically, preclinical
species include mouse, rat, dog, or monkey. The user incorporates compound-specific
physicochemical and \textit{in vitro} data, relevant to the species of interest into the model. The
model prediction is then compared to observed data in the preclinical species. The
focus here is to validate the model, inherent assumptions, and input data used for
predicting the individual PK processes (distribution, clearance and absorption). The IV
profile is predicted, using the observed clearance as input, to assess the predictability of
the distribution. Here compound class and physiochemical properties may influence the
methodology selection, and Kp scalars may be incorporated or comparisons to
allometric scaling methodologies made. Clearance prediction methodologies are
dependent on the clearance mechanism involved, the BDDCS (Biopharmaceutical Drug
Disposition and Classification System) classification, and whether IVIVC across species
had been established. The prediction of absorption is assessed in preclinical species
using an optimized disposition model, with solubility and permeability data as inputs.
The methodology applied will depend on the BCS (Biopharmaceutical Classification
System) classification and whether transporters are involved. Where the preclinical
prediction is reasonably accurate, the user will then move forward with human
prediction. If there are disconnects between PBPK model predictions, based on \textit{in vitro} data, and observed \textit{in vivo} preclinical data, the model assumptions must be critically evaluated. In some cases, additional \textit{in vitro} experiments may be required, to refine critical hypotheses during learn and confirm cycles.

For the human PK prediction, the same physicochemical data are used, in combination with human specific \textit{in vitro} data. Human plasma concentration-time profiles after intravenous (IV) and oral (PO) administration, under single or multiple dose conditions, may then be simulated by the user. This FIH prediction strategy is described in Figure 1.

The accuracy of PBPK predictions of human PK has been investigated in previous publications (Jones \textit{et al.}, 2006, Parrot \textit{et al.}, 2005, De Buck \textit{et al.}, 2007, Jones \textit{et al.}, 2011, Jones \textit{et al.}, 2012, Zhou \textit{et al.}, 2016, Saehang \textit{et al.}, 2018, Naga \textit{et al.}, 2022, Mao \textit{et al.}, 2023). Based on these summaries, volume predictions were found to be generally within 2-fold of the observed. There was a general trend for under-prediction of \textit{in vivo} clearance from \textit{in vitro} metabolism data. Empirical scaling factors, based on \textit{in vitro} clearance ranges in hepatocytes or microsomes, from human or rat, have been proposed to address this well-documented underprediction of clearance (Hallifax and Houston 2012, Wood \textit{et al.}, 2017). In addition, \textit{in vitro} clearance scalars, based on plasma protein binding, have been shown to improve \textit{in vivo} clearance predictions (Jones \textit{et al.}, 2022). Similarly, maximum drug concentration ($C_{\text{max}}$) and area under the plasma concentration-time curve (AUC) after oral administration in human was generally predicted within 2-fold of observed (Naga \textit{et al.}, 2022, Mao \textit{et al.}, 2023). First-order absorption models, where rate ($k_a$) and extent ($f_a$) of absorption were based on averaged preclinical data or \textit{in vitro} permeability data, performed similarly well for four compounds investigated (Mao \textit{et al.}, 2023). Mechanistic predictions, using physiochemical properties including LogD, aqueous and biorelevant solubility, and \textit{in vitro} permeability and predicted observed human $C_{\text{max}}$ and AUC reasonably well (Naga \textit{et al.}, 2022). Mechanistic absorption, predicted based on both measured and predicted physicochemical properties, performed similarly well in this data set.

The application of PBPK during drug discovery has been illustrated through three case studies.
Key Recent Advances

Case Study 1

In this case study, PBPK modelling was used to predict the human PK of a compound following oral absorption. In addition, the effect of CYP3A4 auto-inactivation and drug-drug interaction (DDI) liability as a perpetrator of CYP3A4 inhibition were also predicted. The model and its assumptions were first verified in preclinical species (rat and dog), whereby distribution, clearance and absorption processes were each verified separately.

The PBPK model (Simcyp version 19) was developed using physicochemical data (LogP and pKa), binding data (fu and B:P), solubility and permeability measurements as well as clearance. Oral absorption was described by a mechanistic model (Advanced Dissolution, Absorption and Metabolism (ADAM)). Input parameters specific to the species of interest were used as necessary.

Preclinical model development included incorporation of observed in vivo clearance in preclinical species. The steady state volume of distribution ($V_{SS}$) and plasma concentration profile after IV administration were predicted accurately in preclinical species using the Rodgers and Rowland tissue composition equations (Rodgers and Rowland 2005 and 2007) with a tissue partition coefficient (Kp) scalar of 0.5 (Figure 2). $V_{SS}$ is predicted by combining the predicted volume in blood and the individual tissues which make up the PBPK model. Individual tissue volume is predicted from the specific tissue volume and tissue-plasma partition coefficient (Kp). Three methods are available for prediction of Kp values. The method described by Rodgers and Rowland uses Log P/D and protein binding as input and considers ionization (pKa). In the event of a discrepancy between predicted and observed $V_{SS}$, a global Kp scalar may be applied to simultaneously scale all predicted tissue specific Kp values to an identical extent. The inherent assumption is that the apparent discrepancy is due to a systematic error in the prediction of tissue partitioning, rather than a failure to predict partitioning in a specific tissue.

Absorption in preclinical species was predicted using passive permeability, measured in human colorectal adenocarcinoma (Caco-2) cells, and an intrinsic transcellular...
permeability ($P_{\text{trans},0}$) value was derived from the effective permeability predicted in human ($P_{\text{eff,man}}$) based on this in vitro data. The $P_{\text{trans},0}$ was then applied across species to determine their subsequent permeability predictions. Solubility data obtained in aqueous buffer and simulated intestinal media was used to estimate intrinsic solubility, a solubility factor and bile micelle partitioning in the Simcyp In Vitro analysis (SIVA) Toolkit. The low dose data (0.1 and 1 mg/kg in rat and 1 and 10 mg/kg in dog) were predicted well. Results following oral solution of 1 mg/kg to the rat are shown in Figure 2.

To build confidence further, the $V_{\text{SS}}$ was calculated using the Oie and Tozer method for comparison (Oie and Tozer 1979), which was in close agreement with the Rodgers and Rowland predicted $V_{\text{SS}}$. Clearance was extrapolated from hepatocyte data as well as predicted using single species allometric scaling from preclinical in vivo data. There was a reasonable in vitro to in vivo extrapolation (IVIVE) relationship across species with the predicted clearance being within 2-fold of the observed value.

After model verification using preclinical data, a human PBPK model was developed to predict the PK after oral administration in human subjects. Human PK was predicted over a range of doses (10 – 300 mg QD). Several scenarios (HH $CL_{\text{int}}$, dog-scaled allometric $CL_{\text{IV}}$ and rat-scaled allometric $CL_{\text{IV}}$) were each employed in the model to account for uncertainty in the prediction of human CL. For the HH $CL_{\text{int}}$ scenario, both tablet and solution formulations were simulated to account for uncertainty in the prediction of absorption. Assuming a target exposure of 20 µg/mL at trough, the efficacious dose was predicted to be 100 mg with HH $CL_{\text{int}}$ input and rat-scaled $CL_{\text{IV}}$ input, and 80 mg with dog-scaled $CL_{\text{IV}}$ input. Dose predictions using HH $CL_{\text{int}}$ input for tablet formulation are shown in Figure 3.

Following human dose prediction, the PBPK model was applied to investigate DDI potential. This compound is predicted to be a sensitive substrate of CYP3A4 based on in vitro incubation in recombinant human CYPs. In order to scale the over-expressing recombinant data, inter system extrapolation factors (ISEFs) are applied to correct for differences in catalytic activity per unit enzyme between recombinant and native human liver microsomal CYP enzymes. These factors have been determined internally by comparing parallel incubations with individual CYP isoform probes substrates in
recombinant and pooled human liver microsome incubations. Accounting for inbuilt ISEFs, and individual CYP isoform abundance in the human liver, fm\textsubscript{CYP3A4} was predicted to be \textasciitilde99\%. In addition, this compound is a mechanism-based inactivator of CYP3A4 \textit{in vitro}. Incorporating these properties into the model lead to a prediction of moderate auto-inhibition of CYP3A4. Following repeat 300 mg doses (QD for 14 days), the model predicted an increase in midazolam (sensitive CYP3A4 substrate) exposure, with geometric mean ratios of up to 2.43 and 1.92 for AUC\textsubscript{last} and C\textsubscript{max}, respectively. Given that this compound is a low clearance drug with little hepatic extraction, incorporating CYP3A4 inactivation had little effect on clearance. On the other hand, midazolam is a high clearance drug with high hepatic extraction, hence, inactivation has a greater impact on its overall clearance (Figure 4).

To conclude, a fit-for-purpose preclinical PBPK model was successfully verified by demonstrating accurate simulation of observed plasma concentration data after IV or PO administration in rat or dog. Additional confidence in the model was built through demonstration that the results of other standard approaches for clearance and V\textsubscript{SS} prediction were in agreement with PBPK estimates. However, model verification based on preclinical data is not sufficient justification for the assumption of accurate human prediction. In this case, different scenarios for clearance were simulated in human to account for uncertainty in the prediction of human CL. In addition, a manual sensitivity analysis was completed to account for uncertainty in \textit{the in vitro} CYP3A4 inhibition data. Perpetrator DDI simulations were performed over a 10-fold range of CYP3A4 K\textsubscript{I} to account for uncertainty in free fraction in the \textit{in vitro} incubation or other factors which may contribute to a disconnect between \textit{in vitro} and \textit{in vivo} inhibition potency. In the event of any uncertainty, sensitivity analyses or simulation of alternative scenarios are strongly recommended.

**Case Study 2**

In the second case study, a PBPK model was developed to support candidate selection, from a chemical series, for Phase 1 clinical trials. The chemical series included a lead compound and several compounds with similar properties. All compounds in the series were analogs of an approved drug.
All candidate compounds were neutral with similar chemical structure, low permeability, and high solubility. Therefore, all members of the series were assumed to have similar mechanisms governing their PK properties. The goal of this discovery program was to develop a compound with lower clearance and improved potency, relative to the approved drug, resulting in an improved dosing regimen. Therefore, the primary objective of the PBPK modelling was to identify the most promising candidates, from the available compounds in the chemical series, for further studies in human subjects.

First, a PBPK model was developed for the approved compound (Simcyp version 20). This included input parameters such as physicochemical properties, binding properties, and \textit{in vitro} metabolism data. Clearance was initially extrapolated from human liver microsome (HLM) data. However, in this case clinical data were available for the approved compound. Therefore, data for the approved drug were used for model development and verification. This included plasma concentration data after a single oral dose of 250 mg of the approved compound. An empirical scaling factor, based on Wood \textit{et al.}(2017), was used to account for clearance underprediction from \textit{in vitro} intrinsic clearance (CL\textsubscript{int}) data. In their publication, Wood \textit{et al.} (2017) proposed predicted unbound intrinsic clearance empirical scaling factors for clearance generated in hepatocytes or liver microsomes, from human or rat, based on observed unbound intrinsic clearance. These empirical scaling factors compensate for the trend of underprediction of \textit{in vivo} clearance by metabolism data collected in hepatocytes or liver microsomes. Clearance extrapolation assumed that the scalar applied for the approved drug could be used across the other compounds. Allometry of preclinical data was also used to predict \textit{in vivo} clearance to justify this approach. Overall, reasonable predictions (within 2-fold of observed) were achieved using a full PBPK model (V\textsubscript{SS} predicted with the Rodgers and Rowland method), scaled CL\textsubscript{int} derived from HLM data and a mechanistic oral absorption model incorporating predicted mechanistic effective permeability. As this compound was a high solubility compound, it was treated as a solution formulation within the model. This model development and verification approach relied on clinical data from a similar approved compound. However, the same approach is recommended regardless of the availability of clinical data for similar drugs.
After verification of the PBPK model by simulation of the observed clinical plasma concentration data of the similar approved compound (within 2-fold of observed), the model was applied across all other compounds in the chemical series. This comparison was conducted to support identification of the series compounds with the best chances of success and to aid candidate selection. The corresponding physicochemical properties and *in vitro* inputs were used for each compound.

The target exposure for each compound was derived based on the EC$_{90}$ cell potency data determined *in vitro*. From these, a minimum concentration (C$_{\text{min}}$) target value was determined. Efficacy was assumed to occur only when exposure was greater than C$_{\text{min}}$ for the duration of the dosing interval. Hence, the dose was calculated to achieve a corresponding C$_{\text{min}}$ value at steady-state above the EC$_{90}$. For each compound, dose predictions were simulated for QD (once daily) dosing to improve the dosing regimen compared to the already approved clinical drug. A dose of less than 100 mg QD was deemed as meeting the pre-defined criteria of a lower dose and dosing regimen compared to the clinical comparator. While predicted systemic clearance was similar across all compounds in the chemical series, the unbound intrinsic clearance (CL$_{\text{int, u}}$) was highly variable.

The results of the PBPK work have been summarized in a heatmap, to aid with candidate selection. The heatmap was created based on the lead compound properties, by plotting free potency as a function of CL$_{\text{int, u}}$ (*Figure 5*). The PBPK model was applied using a series of different CL$_{\text{int, u}}$ estimates and dose levels. The shaded region on the heatmap indicates the region of optimal free potency as a function of CL$_{\text{int, u}}$ which is predicted to result in a lower efficacious dose.

As a result of this PBPK modelling, an initial candidate was selected for GLP-toxicity studies and taken into clinical studies.

The major benefit of this modelling approach was its efficiency, supporting rapid candidate selection using predominantly *in vitro* data. Whilst an alternative approach would be the application of allometric scaling, based on *in vivo* PK data, this approach is generally more costly and less time-efficient. Additionally, because CL$_{\text{int, u}}$ is the main driver of the required dose, a PBPK approach is sufficient and more appropriate for the
purpose of candidate selection, given the differences in metabolic mechanisms between preclinical species and humans. Using a bottom-up PBPK approach based on robust \textit{in vitro} data was shown not only to increase efficiency in reaching candidate selection more rapidly, but also to reduce the number of species used in preclinical studies. Overall, PBPK modelling was successfully used in benchmarking against an approved drug to determine optimal properties for candidate selection.

\textbf{Case Study 3}

In this example, the PBPK model was used for FIH dose and DDI predictions for a compound and its active metabolite. The FIH dose prediction target was set to maintain total plasma levels of parent and metabolite equal to or greater than \textit{in vivo} EC\textsubscript{90} values that had been determined in mouse and corrected for plasma protein binding differences.

This compound is a lipophilic diprotic base with moderate solubility and permeability. It is a substrate of P-gp \textit{in vitro}. The compound exhibits moderate binding to plasma. The compound is primarily metabolized by CYP3A4 (fmCYP3A4= 0.90) and forms an active metabolite. The fraction metabolized by CYP3A4 was estimated using \textit{in vitro} phenotyping data generated in recombinant human CYPs and scaled based on appropriate ISEFs and CYP abundance \textit{in vivo}, similar to Case Study 1. Both parent and metabolite were determined to be competitive inhibitors of CYP3A4 in HLMs.

Similar to the previous two case studies, the PBPK modeling approach and assumptions were first verified in preclinical species such as rats, dogs, and monkeys for both parent and metabolite. This verification involved comparing model simulations to \textit{in vivo} data. Validation was conducted separately for each PK process, including distribution and elimination, for both the parent and metabolite.

The PBPK modeling work for each preclinical species was performed using Simcyp Discovery (Version 1), which provides a single interface for a smooth transition of PBPK models from one species to another. The human PBPK model including parent and metabolite was developed in the Simcyp (Simulator version 21), using physicochemical data (Log P, pKa), species-specific binding (fu, B:P), passive permeability and efflux
data from Caco-2 cell lines, in vitro clearance data from microsomes and hepatocytes and non-specific binding data.

The first step of model development was modeling the distribution of parent. In this example, preclinical PK data were available in rats, dogs, and monkeys after IV administration of both parent and metabolite dosed individually (1 mg/kg). Parent and metabolite V_{SS} were predicted using the method proposed by Rodgers and Rowland (2007), in combination with species-specific Kp scalars (Figure 6) optimized to match the observed V_{SS} after IV administration. For human simulations, the Kp scalar was equivalent to the average value of the optimized preclinical species-specific Kp scalars. In addition, allometry was conducted for the prediction of human V_{SS}. The allometry-scaled human V_{SS} estimates, using rat, dog and monkey data after correcting for species-specific fu, were in close agreement with the parent and metabolite V_{SS} values predicted by the Rodgers and Rowland approach.

The parent compound is expected to exhibit significant intestinal loss, mediated by metabolism by CYP3A4 and efflux by P-gp. Absorption modeling for preclinical species was not explored, considering the differences in the extent of intestinal loss across species, which is likely due to cross-species differences in gut wall enzyme and transporter abundance and compound specificity (Chu et al., 2013, Komura et al., 2008).

For predicting clearance of parent and metabolite, in vitro data from liver microsomes and hepatocytes was scaled after correcting for non-specific and blood binding. In vitro CL_{int} estimates were calibrated based on a pre-existing calibration using well studied compounds with diverse properties. The IVIVC was verified first in preclinical species. For CL, there was a reasonable in vitro-in vivo correlation (IVIVC) relationship across species, with the predicted CL within 0.6-1.2-fold of the observed value. The scaled clearance values from LM and hepatocytes were within 2-fold of the observed CL_{IV} in preclinical species for parent and metabolite.

Substrate depletion data for the parent were available in HLM and HH along with formation rate of the metabolite in HLM. There was a disconnect between HLM and HH data with the HLM CL_{int} values being higher. The HLM data were used together to assign the CL_{int} of the parent and the relative formation of the metabolite. Metabolism of
the parent was attributed mainly to CYP3A4, based on \textit{in vitro} phenotyping studies (fm > 0.90). Renal clearance was scaled allometrically, using preclinical species data, and integrated into the PBPK model. For the metabolite, the elimination was described as whole liver metabolic clearance, using a HLM substrate depletion CL\textsubscript{int} value, along with a minor allometry-scaled renal CL. The predicted CL values were also rationalized against hepatocyte scaled and allometry scaled CL estimates.
After verification of the preclinical species PBPK models, a human PBPK model was developed for the prediction of parent and metabolite PK after oral administration. All human simulations were performed using the Simcyp Cancer Population (10 trials of 10 virtual subjects each, 50% female, ages 20 to 95 years). Absorption was described by the mechanistic ADAM model. The passive permeability and P-gp intrinsic clearance were derived by fitting Caco-2 transwell assay data using three-compartment model in SIVA Toolkit (Version 4). Human simulations were performed using solution formulation, due to the high solubility of the parent compound (>100 µg/mL at pH range 1.2 – 7.4).

Efficacious dose was predicted to be 500 mg (QD), based on the assumed therapeutic target of parent and metabolite total plasma concentration greater than unbound \( C_{\text{min}} \) (EC\(_{90}\)) (Figure 7).

The human PBPK model was then used, prospectively, to simulate the DDI potential when administered with midazolam under steady state inhibitor conditions (Table 1). The CYP3A4 enzyme in vitro inhibition constant \( K_i \) values for parent and metabolite were measured in vitro (0.5 and 0.1 µM, respectively). Following repeat 500 mg doses (QD for 10 days) the model predicted an increase in midazolam exposure. Midazolam AUC\(_{0-\text{inf}}\) and \( C_{\text{max}} \) GMRs were 1.05 and 1.01, respectively. These results indicated no DDI potential for parent and active metabolite when administered with sensitive CYP3A4 substrates. A sensitivity analysis was performed to evaluate the impact of CYP3A4 \( K_i \) on the estimated DDI magnitude by reducing the \( K_i \) values by 10-fold (Killford et al., 2021). The objective of this sensitivity analysis was to account for the ‘worst-case’ scenario of potentially inaccurate CYP3A4 enzyme Ki estimates from in vitro experiments and to address uncertainty in estimates of parent and metabolite non-specific microsomal binding.

To summarize, a PBPK model was used to predict FIH dose and DDI potential of a compound and its primary metabolite, following multiple dose administration. The PBPK model was verified in preclinical species before performing human simulations. Knowledge from in vitro and preclinical study findings was integrated mechanistically into the human PBPK model. Sensitivity analysis was performed to address the
uncertainty associated with \textit{in vitro} metabolism and CYP3A4 enzyme inhibition data of both the parent and metabolite.
Current Challenges and Knowledge Gaps

PBPK models used for FIH dose predictions are developed based on a learn and confirm cycle (Figure 1). In our proposed standard practice, model builders initially refine estimated $V_{SS}$, followed by incorporation of clearance and finally absorption parameters. Current challenges and knowledge gaps in these three broad categories are summarized.

Mechanistic $V_{SS}$ prediction depends on *in vitro* and physiochemical data. Predicted $V_{SS}$ is highly sensitive to compound type (i.e., acid, base, or ampholyte), pKa, LogP, and plasma fu. Any uncertainties in these parameters may affect the predicted $V_{SS}$. The choice of tissue composition methodology should be carefully made in the context of the properties of the compound and the assumptions associated with the methodology. In the event of a discrepancy between predicted $V_{SS}$ and the observed plasma concentration profile, a global Kp scalar may be applied to recover the observed $V_{SS}$ in individual preclinical species. In Case Study 2, a single Kp scalar value was able to recover the observed $V_{SS}$ and plasma profiles in both rat and dog. If the Kp scalar is similar across species, an average Kp scalar may be applied for human predictions. However, in cases where the Kp scalar is species-specific, different scenarios for Kp scalar would then be investigated in the human model. $V_{SS}$ prediction using empirical approaches, such as the method proposed by Oie-Tozer and single species scaling with plasma free fraction correction, should also be performed, and compared to the value predicted using the mechanistic approach to help rationalize any choices. The current best practice to predicting $V_{SS}$ value for human is based on the totality of evidence, considering both mechanistic and allometric predictions.

Gaps in clearance prediction remain. Therefore, model assumptions must be clearly defined so that any uncertainties in the human PK prediction, or risks to the program, can be properly understood by the development team. In Case Study 2, assumptions included passive- and perfusion-limited distribution, passive absorption, as well as the assumption that hepatic clearance represents the only clearance mechanism, where *in vitro* hepatocyte data scales directly and accurately to human. Any deviations from these assumptions may result in inaccurate dose predictions. Hence, given the...
uncertainty in the clearance, it is important to investigate different scenarios and provide
a range for the dose predictions. When predicting the DDI effect, it is best practice to
simulate a worst-case scenario. For the dose predictions, simulating lower exposure
could result in dose predictions that are too high.

In Case Study 3, when comparing in vitro CL\textsubscript{int} values from HLM data to HH data, there
was a disconnect observed. Although the underlying mechanism is still unclear,
proposed theories for such types of HLM-HH disconnects include permeability-limited
uptake and cofactor-limited metabolism, leading to overprediction of CL from HLM data
(Williamson et al., 2020; Bapiro et al., 2023). Additionally, CYP3A activity has been
reported to be reduced in cryopreserved HH, compared to HLM, resulting in
underestimation of CYP3A4 CL\textsubscript{int} in HH (Williamson et al., 2020). Given the
uncertainties associated with the scaling of in-vitro data, both HLM and HH data should
be investigated, and allometry-scaled CL should also sometimes be considered.

Finally, challenges remain in the prediction of oral absorption from in vitro or preclinical
data. In some cases, precipitation of drug solution in the gut after gastric emptying may
be overpredicted for low solubility drugs where comprehensive input data are not
available (as was the case in Case Study 2). Model predictions rely on aqueous
equilibrium solubility estimates. In vivo, kinetic solubility can be higher due to excipients
which reduce the potential for precipitation in vivo. In such situations, based on
preclinical oral simulations, human simulations may need to be run as a solution.
Different formulation scenarios can then be simulated for the human model. In Case
Study 2, a solution formulation resulted in the best recovery of the observed plasma
centration profiles across all preclinical species, with the exception of higher dose
levels in dog. However, since these higher dose levels were not considered to be
pharmacologically relevant, suspension and solution with precipitation did not need to
be explored in the human model. Additionally, the intended formulation should be
simulated in human (in this case, tablet) and comparisons should be made to dose
predictions from solution simulations or whichever formulation was validated
preclinically.
Prediction of preclinical bioavailability is much more challenging, specifically when first-pass metabolism and active transport (efflux or uptake) in the gut are involved, which was the situation in Case Study 3. This challenge stems from limited published data describing the significant species-specific differences in gut wall enzymes and transporter abundance, activity and specificity (Chu et al., 2013, Komura et al., 2008). Hypothesis testing is recommended, using the preclinical PBPK models to identify key factors leading to a disconnect between simulated and observed oral exposure, such as solubility, fraction escaping gut metabolism or active efflux or uptake intrinsic clearance.
Perspective on Future Directions

Significant progress has been made towards maximizing the impact of PBPK in the support of the discovery stage of drug development. Many elements of PBPK are quite effective now. For example, passive permeability is typically well predicted from in vitro data or application of the MechPeff model. Distribution volume is generally well predicted across preclinical species. This allows application of global Kp scalars, optimized with preclinical data, in the prediction of human VSS.

However, gaps remain. Clearance is generally underpredicted from in vitro systems. In many cases, proposed in vitro clearance scalars have been applied to successfully predict in vivo clearance (Hallifax and Houston 2012, Wood et al., 2017). To date, these methods have largely been focused on CYP mediated metabolism. Often the impact of uncertainty regarding clearance mechanism may be mitigated through the exploration of a range of clearance methodologies. These scalars are expected to continue to require further refinement, as we learn more about other metabolism mechanisms and how in vitro data for these pathways translate to in vivo clearance. Another area for future improvement is prediction of oral absorption. Aqueous solubility and biorelevant solubility may be used to predict dissolution but consider treating the dose as solution when these data are not available. Open communication between CMC (Chemistry Manufacturing and Controls) and PBPK model builders is recommended to maximize the utility of standard formulation testing data in support of mechanistic prediction of drug absorption. In many cases, small changes to these standard experiments can result in significant improvement of the PBPK model. As animal intestinal transporter data (e.g., abundance/activity along the intestinal tract) become increasingly available, active transport can be incorporated as part of absorption pre-verification in preclinical species to increase confidence in human absorption prediction. Furthermore, preliminary PD prediction linked to the PBPK model (i.e., PBPK-PD) can also be helpful in determining appropriate doses in FIH clinical trials in a more comprehensive manner.

Finally, we have attempted to show through our case studies the importance of performing pre-verification steps in preclinical species, to build confidence in data and assumptions. Increased collaboration between model builders and the designers of
preclinical studies is likely to further improve the impact of PBPK during the discovery stage of drug development.

Conclusions

In conclusion, PBPK is a useful tool in the prediction of human pharmacokinetics from in vitro and preclinical data. Although there continues to be room for further improvement in the prediction of in vivo clearance and absorption parameters, in most cases human PK parameters and plasma concentration profiles are predicted within 2-fold of observed data. Despite any current limitations, this mechanistic approach is superior to predictions based purely on allometry or other empirical relationships. In addition, the mechanistic nature of PBPK supports opportunities for hypothesis testing and guides the design of additional in vitro or preclinical studies to improve understanding of drug properties and how they may impact human PK. The goal of this mini review is to provide additional guidance on the real-world application of PBPK in support of the discovery stage of drug development. We hope we have provided accessible guidance on the typical steps involved in developing and applying a PBPK model during drug discovery.

Acknowledgements

Data Availability Statement
The data that support the findings of this study are available on request from the corresponding author. The data are not publicly available due to privacy or ethical restrictions.

References


Footnotes

This work received no external funding.

Reprint requests should be sent to Ian Templeton; ian.templeton@certara.com; 100 Overlook Center, Suite 101, Princeton, NJ 08540

Figures

Figure 1. Real-world application of PBPK in support of drug discovery.

Figure 2. Simulated and observed concentration-time plots following 1 mg/kg IV (A) and 1 mg/Kg PO (B) in rat.

The solid line represents the mean of the simulated concentration-time profiles. Open circles squares and triangles represent individual observed data points (n=3 Sprague-Dawley rats).

Figure 3. Simulated mean human PO PK profile (tablet) for 50 – 300 mg QD using HH CL\textsubscript{int} as CL input into the model.

*Simulated trial design includes 10 trials of n=10 healthy subjects, aged 20 – 50 with 50% females.*

Figure 4. Simulated midazolam (5 mg) mean plasma concentration-time profiles on Day 12 in the absence (solid line) or presence (dashed line) of 80 mg QD of inhibitor administration in healthy subjects.

*Simulated trial design includes 10 trials of n=10 healthy subjects, aged 20 – 50 with 50% females.*

Figure 5. Heatmap based on free potency versus unbound intrinsic clearance for all compounds.

*Individual compounds in the series summarized as points, relative to the approved compound (black diamond). The green shaded region indicates the optimal region of target potency/clearance resulting in lower dose requirements. The black arrow indicates improvements required for non-optimal candidates.*

Figure 6. Simulated and observed IV concentration-time plots of parent (A) and metabolite (B) after 1 mg/kg dose in the rat using Method 3 predicted Kp and Vss and the measured IV CL as input into the model.

*The solid line represents the mean of the simulated concentration-time profiles. Open circles, squares and triangles represent mean observed data points from a study in n=3 animals.*

Figure 7. Simulated human PO PK profile for Compound ABC and its metabolite at 500 mg QD.

*The solid line and dotted line represent the mean of the simulated plasma concentration versus time profile of parent and metabolite (n= 100). The two straight lines on the plot: the dashed line and dashed-dotted line represents the EC90 values of parent and active metabolite, respectively.*
Tables

Table 1. Simulated geometric mean $\text{AUC}_{0-\text{inf}}$ and $C_{\text{max}}$ values and corresponding GMRs for midazolam in the absence and presence of parent and metabolite in cancer patients.

<table>
<thead>
<tr>
<th></th>
<th>$\text{AUC}_{0-\text{inf}}$ GMR</th>
<th>$C_{\text{max}}$ GMR</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\text{In vitro }K_i$</td>
<td>1.05</td>
<td>1.01</td>
</tr>
<tr>
<td>10-fold lower $K_i$</td>
<td>1.23</td>
<td>1.05</td>
</tr>
</tbody>
</table>

GMR: geometric mean ratio; $C_{\text{max}}$, maximal drug concentration; $\text{AUC}_{0-\text{inf}}$, area under the curve for time zero to infinity.
### Distribution (Vss and Profile Shape) considerations
- Physchem and compound class influence tissue composition model selection.
- Incorporation of Kp scalars.
- Passive versus active transport.
- Comparison with scaling from animals.

### Clearance (Half life and Extraction Ratio) considerations
- Clearance mechanism and BDDCS classification.
- IVIVC across species and use of empirical scaling factors.
- Involvement of transporters.
- Linearity of clearance.
- Comparison with scaling from animals.

### Absorption (Fa, Ka) considerations
- BCS classification.
- Passive versus active permeability.
- Solubility across pH in buffer versus simulated intestinal media.
- Linearity of absorption.
- Formulation type
- Comparison with scaling from animals.

**INPUT data** -> Physchem, compound class, binding, in vitro CLint, solubility, permeability

**Validation step** -> IV and PO data in one preclinical species, CL, Vss, F
Figure 2

A

B

Systemic Concentration (μg/mL)

Time (h)

Systemic Concentration (μg/mL)

Time (h)
Figure 6

A

B

Systemic Concentration (ng/mL) vs. Time (h)

Systemic Concentration (ng/mL) vs. Time (h)