Research Article

Development of a Physiologically Based Gut Absorption Model for Probabilistic Prediction of Environmental Chemical Bioavailability

Hsing-Chieh Lin and Weihsueh A. Chiu
Department of Veterinary Physiology and Pharmacology, College of Veterinary Medicine and Biomedical Sciences, Texas A&M University, College Station, TX, USA

Abstract
Absorption in the gastrointestinal tract is a key factor determining the bioavailability of chemicals after oral exposure but is frequently assumed to have a conservative value of 100% for environmental chemicals, particularly in the context of high-throughput toxicokinetics for in vitro-to-in vivo extrapolation (IVIVE). For pharmaceutical compounds, the physiologically based advanced compartmental absorption and transit (ACAT) model has been used extensively to predict gut absorption but has not generally been applied to environmental chemicals. Here we develop a probabilistic environmental compartmental absorption and transit (PECAT) model, adapting the ACAT model to environmental chemicals. We calibrated the model parameters to human in vivo, ex vivo, and in vitro datasets of drug permeability and fractional absorption by considering two key factors: (1) differences between permeability in Caco-2 cells and in vivo permeability in the jejunum, and (2) differences in in vivo permeability across different gut segments. Incorporating these factors probabilistically, we found that given Caco-2 permeability measurements, predictions of the PECAT model are consistent with the (limited) available gut absorption data for environmental chemicals. However, the substantial chemical-to-chemical variability observed in the calibration data often led to wide probabilistic confidence bounds in the predicted fraction absorbed and resulting steady state blood concentration. Thus, while the PECAT model provides a statistically rigorous, physiologically based approach for incorporating in vitro data on gut absorption into toxicokinetic modeling and IVIVE, it also highlights the need for more accurate in vitro models and data for measuring gut segment-specific in vivo permeability of environmental chemicals.

1 Introduction
Thousands of man-made chemicals are present in commerce, in the environment, and in human biomonitoring samples, but only a small fraction have sufficient toxicity data to conduct human health risk assessments. Over the last decade, there has been increasing emphasis on the use of in vitro high-throughput screening assays to fill this gap, as it is not feasible to conduct traditional in vivo laboratory animal toxicity tests for so many compounds. However, in this paradigm, extrapolation from in vivo animal tests is replaced with extrapolation from in vitro tests, or in vitro-to-in vivo extrapolation (IVIVE). IVIVE has largely been focused on a “reverse toxicokinetics” (RTK) approach, converting in vitro test concentrations to equivalent oral dose levels. In particular, the U.S. Environmental Protection Agency (EPA) has proposed a method of high-throughput toxicokinetics (HTTK) to conduct IVIVE and predict the oral equivalent dose (OED) that results in the internal steady-state plasma concentration (Cs), equivalent to the in vitro bioactive concentration (Pearce et al., 2017). The OED is subsequently compared with the exposure forecasts for risk-based prioritization of chemicals for further analysis (Wetmore et al., 2012; Wambaugh et al., 2015; Bell et al., 2018).

Current implementations of RTK-based IVIVE, however, have focused on key parameters of hepatic metabolism, plasma distribution (free versus bound to proteins), and urinary excretion. The common equation of Cs was derived from a simple pharmacokinetic model (Wilkinson and Shand, 1975) and calculated based

This is an Open Access article distributed under the terms of the Creative Commons Attribution 4.0 International license (http://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution and reproduction in any medium, provided the original work is appropriately cited.
on oral dose rate (Dose, mg/kg/h), unbound fraction of chemical in blood (f_unBlood), glomerular filtration rate (GFR, L/h/kg), liver blood flow (Q_Liver, L/h/kg), and hepatic intrinsic clearance rate for first-order metabolism (CL_int, L/h/kg):

$$C_{as} = \frac{\text{Dose}}{\text{GFR} \cdot f_{\text{unBlood}} + \frac{Q_{\text{Liver}} \cdot f_{\text{unBlood}} \cdot \text{CL}_{\text{int}}}{Q_{\text{Liver}} + f_{\text{unBlood}} \cdot \text{CL}_{\text{int}}}}$$

Thus, most approaches assume that the fraction of the oral intake dose that is absorbed in the gastrointestinal (GI) tract (F_ab) is 100% (Wetmore et al., 2012; Wambaugh et al., 2015; Ring et al., 2017). However, it is known that some environmental chemicals have been found to have a low F_ab, e.g., F_ab for DDT and paraquat are lower than 15% (Chedik et al., 2017). Moreover, Wambaugh et al. (2018) demonstrated that the predictive ability of C_as could be improved if bioavailability could be predicted accurately. Therefore, accurately predicting bioavailability of environmental chemicals is an important issue that needs to be resolved for human health risk assessment.

The situation for environmental contaminants contrasts with that of drug development, where the topic of predicting oral bioavailability has received a substantial amount of attention. F_ab of a drug is mainly determined by permeability and solubility in the gut lumen (Yim et al., 2020). There are two types of permeability measurements: effective permeability (P_eff), which is measured by in vivo studies in humans, and apparent permeability (P_app), which is measured in in vitro cell systems, typically Caco-2 cells. Under the efforts to reduce animal testing, several computational models have been developed and integrated with in vitro studies to examine drug intestinal absorption (Billett et al., 2017; Cho et al., 2014). Two classes of computational models have emerged: one is constructed based on physicochemical descriptors of chemicals, i.e., empirical mathematical models and quantitative structure-activity relationships (QSAR, or quantitative structure-property relationship [QSPR]); the other is based on integrating human physiological features, physicochemical properties of chemicals, and their interaction, i.e., physiologically based pharmacokinetic (PBPK) modeling (Cabrera-Pérez and Pham-Thé, 2018).

An established physiologically based model used to predict drug intestinal absorption is the compartmental absorption and transit (CAT) model, developed by Yu and Amidon (1999). Agoram et al. (2001) expanded on this CAT model to construct the “advanced” CAT (ACAT) model, with consideration of hepatic first-pass metabolism and adding more compartments to commonly describe the dissolution and precipitation processes in the first-pass metabolism and adding more compartments to commonly describe the dissolution and precipitation processes in the plasma and other tissues except GI tract and liver (Hsieh et al., 2021), and so did not include renal clearance or metabolism. Thus, the ACAT model includes separate physiologically based liver and kidney compartments, in addition to a lumped compartment for the rest of the body. Next, in vivo permeability and F_ab data from pharmaceuticals were used to calibrate the ACAT model and choose among three different hypoththesized scaling approaches using in vitro Caco-2 apparent permeability measurements. Finally, we derived a new equation of C_as from our ACAT model for IVIVE. Additional details follow.

2 Materials and methods

An overview of our approach is presented in Figure 1. In brief, we first collected human in vivo and in vitro permeability data and corresponding fractional absorption measurements (F_ab) on drugs and environmental chemicals from the scientific literature. The different permeability measurements were used to quantify how permeability scales in vitro-to-in vivo and across gut segments. We then adapted the ACAT model to environmental chemicals. The primary difference between drugs and environmental chemicals is that drugs are usually in a formulated state and undergo dissolution and precipitation processes in the GI tract, whereas environmental chemicals are generally presumed to present in dissolved form. Therefore, the compartments and movement related to dissolution and precipitation processes were removed in the PECAT model. Furthermore, our previous version of the ACAT model had a simple, non-physiological structure (i.e., two-compartment model) to describe kinetics in the plasma and other tissues except GI tract and liver (Hsieh et al., 2021), and so did not include renal clearance or metabolism. Thus, the PECAT model includes separate physiologically based liver and kidney compartments, in addition to a lumped compartment for the rest of the body. Next, in vivo permeability and F_ab data from pharmaceuticals were used to calibrate the PECAT model and choose among three different hypoththesized scaling approaches using in vitro Caco-2 apparent permeability measurements. Finally, we derived a new equation of C_as from our ACAT model for IVIVE. Additional details follow.

2.1 Permeability and F_ab data

GI tract segment-specific permeability data

Intestinal permeability is region-specific, meaning that the absorption of each chemical differs among the different segments of the GI tract (Sjöberg et al., 2013; Sjögren et al., 2015; Lennernäs, 2014). The data for gut segment-specific permeability

**Abbreviations:** (A)CAT, (advanced) compartmental absorption and transit; CI, confidence interval; CL_int, intrinsic clearance rate; C_app, steady-state plasma concentration; EPA, Environmental Protection Agency; f_unBlood, unbound fraction of chemical in blood; GFR, glomerular filtration rate; GI, gastrointestinal; HTTK, high-throughput toxicokinetics; IVIVC, in vitro to in vivo permeability correction; IVIVE, in vitro-to-in vivo extrapolation; Q_eff, oral equivalent dose; P_app, apparent permeability; PECAT, probabilistic environmental compartmental absorption and transit; P_eff, effective permeability; Q_Liver, liver blood flow; RTK, reverse toxicokinetics
The in vivo permeability (Pef) values of pharmaceuticals were mainly extracted from Lennernäs (1998), Lennernäs (2007), and Varma et al. (2012). Briefly, the values had been measured by a perfusion technique in the proximal region of the human jejunum lumen. The in vitro permeability (Pap) data were from Artursson and Karlsson (1991), Bock et al. (2004), Lentz et al. (2000), and Volpe et al. (2007) and were all acquired by a permeability assay using Caco-2 cells. The percentages of absorption were also extracted from the above-mentioned references and compiled in the dataset.

The study of O’Hagan and Kell (2015) provided an additional list of Caco-2 Pap for marketed drugs. This data provided sufficient data points to construct the distributions of in vitro to in vivo permeability correction (IVIVC) factors. Based on the list of chemicals with in vivo Pef, we extracted the corresponding Pap from O’Hagan and Kell (2015).

For in vitro Pap of environmental chemicals, we collected values from the study conducted by Punt et al. (2022) and Turco et al. (2015) (Tab. S1). The review published by Dahlgren et al. (2015) and Sjögren et al. (2015) (Tab. S1) provided comprehensive data for regional intestinal Pap of drugs, monosaccharaides, and amino acids. These Pap were measured in an Ussing chamber with human tissue, including duodenum, jejunum, ileum, and colon. The type of permeability data collected from Sjögren et al. (2015) is the estimated regional intestinal Pef of drugs. Briefly, the authors used time-course plasma concentration profiles to estimate the rate of absorption by a deconvolution approach, and further incorporated intestinal surface area and dose remaining in the gut lumen to estimate Pef in duodenum, jejunum, and colon.

Drug/environmental chemicals in vivo and in vitro permeability and Fabs in human

We located data from 80 drugs and environmental chemicals with both permeability and Fab in humans from the literature (Tab. S2). The in vivo Pef values of pharmaceuticals were mainly extracted from Lennernäs (1998), Lennernäs (2007), and Varma et al. (2012). Briefly, the values had been measured by a perfusion technique in the proximal region of the human jejunum lumen. The in vitro Pap data were from Artursson and Karlsson (1991), Bock et al. (2004), Lentz et al. (2000), and Volpe et al. (2007) and were all acquired by a permeability assay using Caco-2 cells. The percentages of absorption were also extracted from the above-mentioned references and compiled in the dataset.

The study of O’Hagan and Kell (2015) provided an additional list of Caco-2 Pap for marketed drugs. This data provided sufficient data points to construct the distributions of in vitro to in vivo permeability correction (IVIVC) factors. Based on the list of chemicals with in vivo Pef, we extracted the corresponding Pap from O’Hagan and Kell (2015).
al. (2011) in which the $P_{\text{app}}$ was measured by Caco-2 permeability test. The actual values were obtained by GetData Graph Digitizer (version 2.26). The fractions absorbed of environmental chemicals were manually searched and extracted from different literature. The reference sources are listed in Table S2.

We further reviewed the literature cited by Punt et al. (2017) and found a list of publications about relationships of drug $P_{\text{app}}$ values obtained using Caco-2 cells with human absorption, possessing a wider range of $P_{\text{app}}$; thus, we collected the data from these papers (Marino et al., 2005; Mattson et al., 2005; Miret et al., 2004; Turco et al., 2011) and integrated it with another drug database published by Paixão et al. (2010) (source of Caco-2 $P_{\text{app}}$) and Paixão et al. (2012) (source of $F_{\text{abs}}$), creating a supplementary dataset to further validate model performance (Tab. S4).

### 2.2 Physiologically based pharmacokinetic absorption model adaption

An open-source version of the PBPK absorption model, i.e., the ACAT model, for predicting the bioavailability of pharmaceuticals was developed and published by Hsieh et al. (2021). This model describes the process of drug dissolution, precipitation, and absorption in the gut lumen, and considers the physicochemical properties of chemicals and physiological factors of humans. We adapted this model to construct a PBPK absorption model specifically for environmental chemicals, and its structure is presented in Figure 2.

First, we assumed that environmental chemicals are usually exposed in dissolved form in matrixes; therefore, the dissolution and precipitation processes of chemicals and the compartments of undissolved chemical in gut lumen segments were removed. To develop a new approach for IVIVE of reverse dosimetry, we replaced the central and peripheral compartments with a traditional PBPK model, including the compartments liver, kidney, and the rest of the body. The details of the updated model equations are compiled in the supplementary material, Section 1.

### 2.3 Model parameterization

#### Updated PBPK absorption model parameters

Most of the physiologically specific and chemical-related parameters were taken from the same literature as in the study by Hsieh et al. (2021). The additional parameters are mainly related to adding the compartments for kidney and the rest of the body. Additional physiological parameters included fractions of cardiac output flow, body weight, and microsomal proteins, while the chemical-related parameters included fractions unbound in organs, partition coefficients, and metabolism-related rates and coefficients. The detailed parametric information is listed in Tables S5 and S6.

A sensitivity analysis was conducted to verify that the structural changes in terms of fixing chemical-related parameters did not significantly affect the prediction of $F_{\text{abs}}$ (details presented in the supplementary material, Section 2).

#### Segment-specific permeability scaling ratios

To consider the segment-specific absorption in the gut, generic distributions of regional intestinal permeability ratios were proposed. These ratios only consider the segment-specific permea-
bility of the gut (already adjusted for physiological characteristics such as surface area) and do not include other physiological characteristics of the gut that are already included in the model structure. Based on the data listed in Table S1, segment-specific permeability scaling ratios from segment \( i \) relative to jejunum or colon permeability were calculated (\( P_{i}P_{\text{jej}} \) and \( P_{i}P_{\text{col}} \), respectively). The reason for these two methods is that the \textit{in vivo} \( P_{\text{eff}} \) and \textit{in vitro} \( P_{\text{app}} \) are measured in the proximal region of the human jejunum lumen and in Caco-2 cells from human colon tissue, respectively. The Shapiro-Wilk normality test with original and log-transformed values was used to investigate whether the ratios fit a normal or lognormal distribution. The correlation matrix for Pearson’s correlation coefficient (R) was also examined to evaluate dependencies between these ratios. For prospective prediction, Monte Carlo (MC) sampling from these ratios was performed using a multivariate normal distribution to capture their correlation structure.

\textit{In vitro-to-in vivo permeability correction (IVIVC) factors}

It has been shown that there can be discrepancies between \textit{in vivo} and \textit{in vitro} measurements when using \textit{in vitro} apparent permeability (\( P_{\text{app}} \)) as a parameter input (Sun et al., 2002). Therefore, \textit{in vitro-to-in vivo} correction (IVIVC) factors have been proposed as an empirical approach to address this issue. In this study, IVIVC factors were calculated by dividing the \textit{in vivo} \( P_{\text{eff}} \) by the \textit{in vitro} \( P_{\text{app}} \) when both measurements were available (Tab. S2, S4). If more than one \( P_{\text{eff}} \) was available for a given chemical, the average of the \( P_{\text{eff}} \) values was used to calculate the ratio. As with gut segment ratios, the Shapiro-Wilk normality test was used with original and log-transformed values to assess whether the ratios fit a normal or lognormal distribution, and MC simulation was used for prospective prediction.

2.4 Model evaluation: Predictions of \( F_{\text{abs}} \) based on different extrapolation assumptions for permeability

\textit{Permeability extrapolation assumptions}

In predicting the \( F_{\text{abs}} \) based on \textit{in vitro} Caco-2 permeability measurements, we assumed that absorption takes place only via passive transport and neglected intestinal active transport and intestinal epithelial metabolism. Only absorption in the small intestine and large intestine were considered (i.e., no absorption from the stomach). Because the exposure to environmental chemicals is assumed to be chronic, we simulated with continuous exposure for 10,000 hours to predict the \( F_{\text{abs}} \) under steady-state conditions. The \textit{in vitro} permeabilities we tested ranged from \( 10^{-9} \) to \( 10^{-2} \) cm/s.

We evaluated three different assumptions for the relationship between \textit{in vitro} Caco-2 permeability and \textit{in vivo} segment-specific permeability:

1. Assuming that \textit{in vitro} \( P_{\text{app}} \) is equal to \textit{in vivo} “jejunum” permeability, then use the distributions of \( P_{i}P_{\text{j}} \) ratios to scale to segment-specific permeabilities. This is the typical assumption commonly used in the literature.

2. Assuming that \textit{in vitro} \( P_{\text{app}} \) is directly equal to \textit{in vivo} “colon” permeability, then use the distributions of \( P_{i}P_{\text{c}} \) ratios to scale to segment-specific permeabilities. This assumption tests the hypothesis that because Caco-2 cells are derived from the colon, they are more representative of colon permeability.

3) First convert \textit{in vitro} \( P_{\text{app}} \) to \textit{in vivo} \( P_{\text{eff}} \) in the “jejunum” by using the IVIVC factors, then use the distributions of \( P_{i}P_{\text{j}} \) ratios to scale to segment-specific permeabilities. This assumption uses a two-step extrapolation approach.

The uncertainty distributions for the predictions were derived by using the distributions of segment-specific ratios and IVIVC factors as inputs and MC simulation with 10,000 samples.

\textit{Model evaluation}

Although the simulation was executed under the assumption of only considering passive transport, the issues of active transport and metabolism still should be discussed. The Biopharmaceutics Drug Disposition Classification System (BDDCS), based on solubility and extent of metabolism, can be used to classify drugs and provide information on the effect of active transport on the absorptive process in the gut. We applied the classification criteria to both drugs and environmental chemicals (Benet et al., 2011; Bocci et al., 2022; Hosey et al., 2016; Wu and Benet, 2005; Tab. S2-S4). The detailed characteristics of each BDDCS class are compiled in Table 1. The subset of drugs with \textit{in vivo} \( P_{\text{eff}} \) and corresponding \( F_{\text{abs}} \) was used to evaluate the predictive ability of the PECAT model. Specifically, the \( F_{\text{abs}} \) distributions predicted by PECAT were compared with the corresponding observed \( F_{\text{abs}} \). Mean absolute error (MAE) compared to observations was then calculated to evaluate the model’s precision of median prediction:

\[
\text{MAE} = \frac{\sum \left| F_{\text{abs,pred}}^{\text{50}} - F_{\text{abs,obs}} \right|}{N} = \frac{\sum |\Delta F_{\text{abs}}|}{N}
\]

where \( \Delta F_{\text{abs}} \) is the difference between the simulation median and observed values. We also calculated the mean of \( \Delta F_{\text{abs}} \) (mean errors; MEs) to examine the bias. MAEs and MEs were calculated separately for different BDDCS classes. We also utilized the percentage of the number of chemicals predicted within ±20% of \( \Delta F_{\text{abs}} \) as an index to evaluate predictive accuracy (Paixão et al., 2012). The best performing extrapolation assumption was used for subsequent calculations.

2.5 Updated equation of \( C_{\text{ss}} \) derived from PECAT model

Based on our PECAT model, we derived an updated equation of the steady state concentration \( C_{\text{ss}} \), incorporating \( F_{\text{abs}} \) and extraction ratio (\( E \)) for first-pass metabolism into the term for the oral dose rate, for performing reverse dosimetry for IVIVE:

\[
E = \frac{f_{uBlood}\cdot CL_{\text{int}}}{Q_{\text{Liver}} + f_{uBlood}\cdot CL_{\text{int}}}
\]

\[
C_{\text{ss}} = \frac{\text{Dose} \cdot F_{\text{abs}} \cdot (1-E)}{GFR \cdot f_{uBlood} + Q_{\text{Liver}} \cdot E}
\]
with the concentrations in blood given out by the PECAT model when simulating the relationship of permeability and $F_{\text{abs}}$ under long-term exposure. Figure S7\(^1\) shows that the updated equation of $C_{\text{ss}}$ can successfully predict the concentrations in blood at steady-state using the PECAT model.

### 2.6 Computational software

PECAT modeling with a Monte Carlo algorithm for prediction of $F_{\text{abs}}$ was conducted by GNU MCSim v6.1.0 executed in R environment. Other statistical analysis and visualization of results were performed in R (version 4.1.2) with RStudio as an integrated platform. All raw data and computational codes for modeling and data analysis are available in the supplementary material\(^1\) and on GitHub\(^2\).

### 3 Results

#### 3.1 Segment-specific permeability scaling ratios

The results of distributions of $P_i^{\text{jejumum}}$ and $P_i^{\text{colon}}$ ratios are shown in Figure 3 with summary statistics in Table 2. The average values of the $P_i^{\text{jejumum}}$ ratio of duodenum, ileum, and colon are smaller than 1, meaning that the jejunal permeability is generally greater than that of other segments and the jejunum has better absorption activity (Fig. 3A; Tab. 2), but distributions for all three ratios include 1. Figure 3B shows that distributions of the $P_i^{\text{jejumum}}$ ratio of duodenum, jejunum, and ileum all have mean values greater than 1 (Tab. 2), so that the colon permeability is on average lower than that of other segments, though again the distributions overlap with 1.

Based on the results of normality tests, both kinds of regional intestinal permeability ratios are consistent with being lognormally distributed ($p$-value > 0.05; Tab. 2). As for understanding the relationship among ratios, the correlation matrixes of log-transformed $P_i^{\text{jejumum}}$ ratios shown in Figure S8A\(^1\) suggest that the relationships of the ratios have relatively low correlation, with an absolute Pearson’s $R$ of less than 0.7 ($p$-value > 0.05). Figure S8B\(^1\) shows that Pearson’s $R$ of $P_i^{\text{jejumum}}$ to $P_i^{\text{duodenum}}$: $P_i^{\text{colon}}$ and to $P_i^{\text{jejumum}}$ colon to $P_i^{\text{duodenum}}$: $P_i^{\text{colon}}$ ratios are 0.80 and 0.82 ($p$-value < 0.05), respectively. This means that they have highly positive correlation; therefore, multivariate normal distributions should preserve the correlation structure of colon ratios when implementing the MC algorithm.

#### 3.2 Distributions of in vitro to in vivo permeability ratio

Figure 4 presents the distribution of IVIVC factors, and the summary statistics are summarized in Table 3. The mean of the IVIVC factor is 0.134, which indicates that, on average, the in vitro apparent permeability is around ten times lower than the in vivo effective permeability. Similar to the segment-specific ratios, the IVIVC factor is consistent with being lognormally distributed ($p$-value = 0.644; Tab. 3), and this distribution was used when performing MC simulations.

#### 3.3 Comparisons of simulated and observed relationship of permeability and $F_{\text{abs}}$

Model validation was conducted for each Scenario. Scenario I incorporated the distributions of the $P_i^{\text{jejumum}}$ ratio, and the comparison of the simulated relationship of permeability and $F_{\text{abs}}$ with the observed in vivo $P_{\text{app}}$ and human $F_{\text{abs}}$ had a MAE of 8.9% (Fig. 5A, D; Tab. 4). Simulating with the same $P_i^{\text{jejumum}}$ ratio distributions but comparing with in vitro $P_{\text{app}}$ of drugs and environmental chemicals, the results in Figure 5A show that the simulated relationship with 95% confidence interval (CI) only captured less than two-thirds of the data points (62%; 49 out of the sum of 70 drugs and 9 environmental chemicals). Additionally, the MAEs of drugs and environmental chemicals were both around 29%, and the MEs of drugs and environmental chemicals were -28% and -29%, respectively, showing an underestimation of $F_{\text{abs}}$ under Scenario I (Fig. 5D; Tab. 4). Under this set of extrapolation assumptions, the PECAT model can predict 41% of drugs (29 out of 70 data points) and 56% of environmental chemicals (5 out of 9 data points) within ±20% error ranges. These results all indicated that directly using $P_{\text{app}}$ and $P_i^{\text{jejumum}}$ ratio as input for the PECAT model could not predict $F_{\text{abs}}$ accurately.

Under Scenario II, we assumed that the input permeability is equal to in vivo $P_{\text{eff}}$ in the colon and incorporated the distributions of $P_i^{\text{jejumum}}$ colon to predict the $F_{\text{abs}}$. As shown in Figure 5B, the simulated relationship curve with 95% CI can capture more than three quarters of the data points (76%; 60 out of 79 data points), i.e., more than Scenario I. Figure 5E shows the difference between median predicted and observed $F_{\text{abs}}$ and the MAEs of drugs and environmental chemicals improved to around 16% and 19%, respectively (Tab. 4). Also, median $F_{\text{abs}}$, for 70% of drugs and 78% of environmental chemicals were within ±20% error ranges (Fig. 5E). The MEs of drugs and environmental chemi-

---

**Table 1: Detailed characteristics of the BDDCS classes referred to by Wu and Benet (2005) and Benet et al. (2011)**

<table>
<thead>
<tr>
<th>BDDCS</th>
<th>Solubility</th>
<th>Permeability</th>
<th>Extent of metabolism</th>
<th>Transporter effects in gut</th>
</tr>
</thead>
<tbody>
<tr>
<td>I</td>
<td>High</td>
<td>High</td>
<td>Extensive</td>
<td>Minimal effects</td>
</tr>
<tr>
<td>II</td>
<td>Low</td>
<td>High</td>
<td>Extensive</td>
<td>Predominant efflux transporter effects</td>
</tr>
<tr>
<td>III</td>
<td>High</td>
<td>Low</td>
<td>Poor</td>
<td>Predominant absorptive transporter effects</td>
</tr>
<tr>
<td>IV</td>
<td>Low</td>
<td>Low</td>
<td>Poor</td>
<td>Both absorptive and efflux transporters could be important</td>
</tr>
</tbody>
</table>

\(^{2}\) https://github.com/hsingchiehlin/PECAT_model
with in vitro $P_{\text{app}}$ between $10^{-7}$ and $5 \times 10^{-6}$ cm/s had a very wide 95% CI of predicted $F_{\text{abs}}$, in which the difference between the upper and lower bound was over 70%. In this interval of permeability, the maximum 95% CI of $F_{\text{abs}}$ ranged from 3% to 99%.

The MAEs of drugs and environmental chemicals were around 13% and 14%, respectively (Fig. 5F; Tab. 4). The PECAT model with this scenario results in the $F_{\text{abs}}$ for 81% of drugs and 78% of environmental chemicals predicted within ± 20% (Fig. 5F). It is noted that the slightly positive MEs of drugs (5%) and environmental chemicals (7%) indicate that the PECAT model under Scenario III may tend to slightly overestimate $F_{\text{abs}}$ (Tab. 4).

The final extrapolation assumption (Scenario III) integrated $P_{\text{app}}$ with distributions of IVIVC factors as well as $P_{i:j}$ ratios to predict $F_{\text{abs}}$. The 95% CI of the simulated relationship curve captured most data points for both drugs (92%; 65 out of 70 data points) and environmental chemicals (89%; 8 out of 9 data points) (Fig. 5C). However, we found that the chemicals, being around -12% and -9% (Tab. 4), respectively, showed that the underestimation in Scenario I was somewhat mitigated under Scenario II. These results overall indicated that the PECAT model, using $P_{\text{app}}$ and $P_{i:j}$ ratio, predicted $F_{\text{abs}}$ better than the simulations under Scenario I.

The data of segment-specific permeability, collected from Dahlgren et al. (2015) and Sjögren et al. (2015), includes permeability measurement from 27 drugs.

**Tab. 2: Summary of distribution of gut segment-specific permeability ratios**

<table>
<thead>
<tr>
<th>Ratio</th>
<th>Mean</th>
<th>SD</th>
<th>Geo Mean</th>
<th>Geo SD</th>
<th>$p$-value</th>
<th>Selected distribution</th>
</tr>
</thead>
<tbody>
<tr>
<td>$P_{i:j}$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Duodenum</td>
<td>0.676</td>
<td>0.303</td>
<td>0.609</td>
<td>1.672</td>
<td>0.746</td>
<td>LogNormal</td>
</tr>
<tr>
<td>Ileum</td>
<td>0.913</td>
<td>0.737</td>
<td>0.678</td>
<td>2.239</td>
<td>0.025</td>
<td>LogNormal</td>
</tr>
<tr>
<td>Colon</td>
<td>0.727</td>
<td>0.846</td>
<td>0.327</td>
<td>4.633</td>
<td>&lt; 0.001</td>
<td>LogNormal</td>
</tr>
<tr>
<td>$P_{i:c}$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Duodenum</td>
<td>1.578</td>
<td>1.123</td>
<td>1.272</td>
<td>2.045</td>
<td>0.114</td>
<td>LogNormal</td>
</tr>
<tr>
<td>Jejunum</td>
<td>11.030</td>
<td>23.239</td>
<td>3.107</td>
<td>4.615</td>
<td>&lt; 0.001</td>
<td>LogNormal</td>
</tr>
<tr>
<td>Ileum</td>
<td>5.459</td>
<td>8.673</td>
<td>2.410</td>
<td>3.440</td>
<td>&lt; 0.001</td>
<td>0.269 LogNormal</td>
</tr>
</tbody>
</table>
Comparing among the three scenarios, we found that the smallest MAEs and the largest percentage of accuracy (within ± 20%) were reached under the extrapolation assumptions in Scenario III, where the IVIVC factors and the $P_i:P_{jejum}$ ratios are both integrated. These results also suggest that the \textit{in vitro} $P_{appl}$ in Caco-2 cells cannot be directly assumed to be equal to \textit{in vivo} $P_{eff}$ in the jejunum.

In further validation with the supplementary database (Tab. S4) under Scenario III (Fig. S9), the simulated relationship curve with 95% CI using the PECAT model with the same two scaling ratios captured 91% of data points (148 out of 162 points), and had overall MAE and ME of 15% and 5%, respectively, as well as 72% of data points (117 out of 163) within a range of ± 20% of predictions. These results are similar to the predictive performance found using the main dataset. These results confirmed that the PECAT model using $P_{appl}$ as input and incorporating IVIVC factors and $P_i:P_{jejum}$ ratios can be used to predict the range of possible $F_{abs}$ values for environmental chemicals.

To examine possible differences in performance across chemical classes, we used BDDCS classes. Table 4 shows that MAEs of drugs in Classes III and IV were generally higher than in Classes I and II, no matter which scenario was implemented. The smaller MAEs of Class II (Tab. 4) could suggest that $F_{abs}$ of chemicals that are dominantly affected by efflux transporters can still be predicted well using Caco-2 data and our PECAT model with the assumption of only considering passive absorption. The results of MEs demonstrated that, in Scenarios I and II, the $F_{abs}$ in all classes were generally underestimated, but in Scenario III, only $F_{abs}$ of Class III were on average overestimated (Tab. 4). The validation results with the supplementary database led to the same conclusion (Fig. S9). Based on the BDDCS classification criteria, we found that most environmental chemicals used to assess performance were similar to Class II and only one was similar to Class IV (Tab. S3). The ME of environmental chemicals in Class II also demonstrated results that were similar to those for drugs, but the MAE of environmental chemicals in Class II was higher than for drugs in the same class.

4 Discussion

Since the landmark National Academies report \textit{Toxicity Testing in the 21st Century: A Vision and Strategy} (NRC, 2007), it has been recognized that transitioning to greater reliance on \textit{in vitro} testing strategies requires the development of IVIVE approaches to convert \textit{in vitro} bioactive concentrations to equivalent exposure levels that can be used for human health risk characterization. The development of RTK-based IVIVE over the last 10 years (Rotroff et al., 2010; Wambaugh et al., 2014, 2015, 2018; Wetmore et al., 2012, 2015) has focused on developing...
methods for quantifying distribution (plasma protein binding assays) and metabolism (e.g., hepatocyte suspensions), relying on simplifying assumptions for other aspects of toxicokinetics – absorption (e.g., 100% gut absorption) and excretion (passive glomerular filtration for free chemical). In particular, while oral absorption is recognized as a major issue for drugs, the advances in this area from pharmaceutical development have not been applied to environmental chemicals. Here, we aim to address this gap by adapting approaches used for pharmaceuticals, focusing on integrating the most commonly used drug models, i.e., in vitro Caco-2 permeability and the in silico ACAT model, to develop our PECAT model.

Based on our literature review, we found a high degree of chemical-to-chemical variability in the relationship between in vitro permeability and in vivo fraction absorbed. This variability has two sources: differences between in vitro apparent permeability and in vivo effective permeability, and differences in in vivo permeability across gut segments. These differences are no doubt due to numerous physicochemical factors as well as differences in the degree of active versus passive transport. For example, cases when the IVIVC is far away from 1 are likely due to the different expression levels of certain active uptake/efflux transporters and/or different paracellular permeability between Caco-2 cells and human jejunum (Vaessen et al., 2017; Youhanna and Lauschke, 2021). The different physiology and proteomes of Caco-2 cells and human jejunum implied applicability domain limitations when applying our PECAT model to predict bioavailability of compounds transported mainly via active transcellular and paracellular pathways (Youhanna and Lauschke, 2021). However, although data on such factors may be available for individual pharmaceuticals to enable adjustments to be made, there is much less such information for environmental chemicals, so we decided to take an empirical approach and treat this chemical-to-chemical variability, as well as inaccuracies due to lack of this information, as an uncertainty for any particular substance.

As for the choice of the empirical approach, Sun et al. (2002) used a linear regression model to describe the relationship of $P_{eff}$ and $P_{app}$, and this method is equivalent to our ratio method but with the regression slope assumed to be fixed at 1. To assess the impact of this assumption, we also conducted linear regression modeling without fixing the slope (supplementary materials, Section 3.1), which demonstrated that fitting the slope did not result in any noticeable improvements. Because the ratio approach is simpler, we retained that in our analysis. Moreover, there were sufficient data from pharmaceuticals on both sources of variability to develop probabilistic distributions that could be integrated into the overall modeling framework. Indeed, we found that utilizing these uncertainty distributions enabled the PECAT to be accurate with a MAE of 13%. Thus, the PECAT model, along with the corresponding updated $C_{ss}$ equation, can be readily incorporated into RTK-based IVIVE, enabling more accurate risk prioritization.

<table>
<thead>
<tr>
<th>Scenario</th>
<th>Permeability</th>
<th>BDDCS</th>
<th>Overall</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Class I</td>
<td>Class II</td>
<td>Class III</td>
</tr>
<tr>
<td>Mean absolute errors</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>I</td>
<td>Drug $P_{eff}$</td>
<td>5.13</td>
<td>2.99</td>
</tr>
<tr>
<td></td>
<td>Drug $P_{app}$</td>
<td>29.30</td>
<td>22.65</td>
</tr>
<tr>
<td></td>
<td>Envi. $P_{app}$</td>
<td>-</td>
<td>24.36</td>
</tr>
<tr>
<td>II</td>
<td>Drug $P_{app}$</td>
<td>9.01</td>
<td>7.77</td>
</tr>
<tr>
<td></td>
<td>Envi. $P_{app}$</td>
<td>-</td>
<td>15.82</td>
</tr>
<tr>
<td>III</td>
<td>Drug $P_{app}$</td>
<td>5.10</td>
<td>8.99</td>
</tr>
<tr>
<td></td>
<td>Envi. $P_{app}$</td>
<td>-</td>
<td>14.41</td>
</tr>
<tr>
<td>Mean errors</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>I</td>
<td>Drug $P_{eff}$</td>
<td>4.97</td>
<td>2.95</td>
</tr>
<tr>
<td></td>
<td>Drug $P_{app}$</td>
<td>-29.30</td>
<td>-21.91</td>
</tr>
<tr>
<td></td>
<td>Envi. $P_{app}$</td>
<td>-</td>
<td>-24.36</td>
</tr>
<tr>
<td>II</td>
<td>Drug $P_{app}$</td>
<td>-7.81</td>
<td>-1.55</td>
</tr>
<tr>
<td></td>
<td>Envi. $P_{app}$</td>
<td>-</td>
<td>-5.02</td>
</tr>
<tr>
<td>III</td>
<td>Drug $P_{app}$</td>
<td>3.61</td>
<td>8.85</td>
</tr>
<tr>
<td></td>
<td>Envi. $P_{app}$</td>
<td>-</td>
<td>7.45</td>
</tr>
</tbody>
</table>

\(^a\) Only one drug or environmental chemical in the group.
A. Scenario I

Percentage of absorption

$P_{\text{eff}}$ or $P_{\text{app}}$ = Jejunum permeability (cm/s)

B. Scenario II

Percentage of absorption

$P_{\text{app}}$ = Colon permeability (cm/s)

C. Scenario III

Percentage of absorption

$P_{\text{app}} \Rightarrow P_{\text{eff}}$ = Jejunum permeability (cm/s)
Fig. 5: Comparisons of observed and predicted $F_{\text{abs}}$ as a function of the permeability (cm/s) under different scenarios

(A) $P_{\text{eff}}$ or $P_{\text{app}}$ assumed to be equal to jejunum permeability and incorporated with the distribution of $P_i:P_{\text{jejunum}}$, (B) $P_{\text{app}}$ assumed to be equal to colon permeability and incorporated with the distribution of $P_i:P_{\text{colon}}$, and (C) $P_{\text{eff}}$ converted from $P_{\text{app}}$ by using IVIVC factor and assumed to be equal to jejunum permeability, thereby incorporating with the distribution of $P_i:P_{\text{jejunum}}$. The solid lines represent the median of predicted $F_{\text{abs}}$, and shaded areas are 95% confidence interval. (D-F) Median differences between the observed and predicted percentage of absorption ($\Delta F_{\text{abs}}$) with 95% CI. The in vivo data points of pharmaceutical chemicals were collected from Lennernäs (1998 and 2007) and Varma et al. (2012) and the in vitro ones were compiled from Artursson and Karlsson (1991), Bock et al. (2004), Lentz et al. (2000) and Volpe et al. (2007). The in vitro data of environmental chemicals was from Punt et al. (2022) and Turco et al. (2011).
When using Caco-2 permeability data to estimate in vivo permeability, the effect of various experimental settings such as pH value and absence/presence of transporter inhibitors on the IVIVC ratio should be discussed. We summarized the detailed information on experimental settings for the chemicals used to derive our IVIVC ratios in Tables S2 and S4, and examined the distributions based on different pH value and absence/presence of transporter inhibitor. These results are summarized in the supplementary materials, Section 3.2; we found that there are also no significant differences between distributions considering all IVIVC ratios or only considering IVIVC ratios under specific experimental settings, meaning that different pH value and presence of efflux transporter inhibitor would not significantly affect the distribution of IVIVC ratios for the datasets we utilized. We also should note that our empirical approach for IVIVC and segment-specific permeability ratios incorporates uncertainties related to passive vs. active transport. Specifically, while our model assumes only passive transport, the in vivo effective permeability and ex vivo apparent permeability data used to build these distributions include both active and passive transport. Thus, the uncertainty distributions for the IVIVC and segment-specific permeability ratios include the errors from assuming only passive transport. That being said, it is likely that our model is more suitable for chemicals absorbed by passive processes.

Our approach has a number of limitations that require additional research. Most importantly, our model assumes that gut absorption is due to passive diffusion only. Although the model structure includes the possibility of active transport (both influx and efflux) and intestinal metabolism, these features are “turned off” (parameter values set to zero) in the simulations presented here because Caco-2 apparent permeability measurements (in the absence of inhibitors) do not provide any separate information on these processes. Thus, chemical-specific data from newer in vitro models (Youhanna and Lauschke, 2021) offer the possibility of improving the accuracy of PECAT model predictions. However, in the meantime, we believe the probabilistic outputs from PECAT appropriately reflect these uncertainties, though the accuracy appears different for different chemical classes.

In particular, although Caco-2 apparent permeability measured without transporter inhibitors cannot provide detailed information on absorptive and metabolic processes, BDDCS can assist in generally understanding the effects of active efflux and absorptive transporters on the performance of our model. In Figure 5, most data points of classes III and IV are located far away from the prediction line with more than ± 20% differences. Based on Table 1, the absorption process of drugs in these two classes is determined by absorptive transporters and both absorptive and efflux transporters (Wu and Benet, 2005; Benet et al., 2011). This suggests that adding information on active transporters should increase accuracy when predicting the $F_{abs}$ of chemicals in Classes III and IV. The negative mean error of drugs in Class II (under Scenario III) can also explain that the assumption of “switched-off” active transporters results in overpredicted $F_{abs}$, because the absorption of Class II drugs is affected by efflux transporters. On the other hand, O’Hagan and Kell (2015) compared the apparent permeability measured without and with efflux inhibitors, and they concluded that there are no order-of-magnitude discrepancies between the two kinds of permeability. This is consistent with our finding that MAEs of Class II drugs are similar to those of Class I drugs, meaning $F_{abs}$ of Class II drugs can be well predicted by our PECAT model.

The second key limitation is that we found very few environmental chemicals with both Caco-2 in vitro data and in vivo measurements of gut absorption. Although Caco-2 data can be generated relatively quickly, it is likely that in vivo measurements will need to rely largely on historical data, and systematic evidence mapping may be useful to determine the chemical space for which such data are available. In addition, we tried to use BDDCS criteria to classify environmental chemicals. However, the limited data on metabolism and solubility (Hosey et al., 2016; Wu and Benet, 2005) means that logP alone is used to classify metabolism, and ExpoCast with water solubility data from the EPA CompTox Chemical Dashboard is used to obtain specific solubility (dose number). These limitations could result in significant misclassification. Therefore, we also referred to other literature to classify the environmental chemicals. Most environmental chemicals in our study were classified as Class II (details shown in Tab. S3). Nonetheless, the assessment of PECAT performance by class is challenging and limited. Therefore, the PECAT model as currently formulated can, depending on the measured Caco-2 apparent permeability, lead to a very large range of predictions for permeability and fraction absorbed. At this point, it may be appropriate to view the current PECAT model as an initial screening tool in a tiered approach for environmental chemicals.

Thus, it will be important to test the PECAT model with additional datasets to better characterize its performance, particularly since the input distributions for IVIVC factors and gut segment-specific permeability ratios are based on drugs, which generally inhabit a different space of physicochemical properties as compared to environmental chemicals.

Finally, chemical exposures occur via a variety of environmental media, and absorption after oral exposure, in particular, may be dependent on whether the compound is found in water, food, or soil, as well as whether the individual is in a fasted state. The PECAT model does not currently address such media-specific and time-dependent differences in absorption. However, for environmental chemicals, it is likely difficult to obtain data to parameterize these processes. The environmental chemical-specific Caco-2 permeability dataset is still limited, and development of a large dataset of uniformly conducted experiments would be helpful to improve application and testing of the PECAT model for environmental chemicals.

Overall, we have developed a probabilistic, physiologically based modeling framework to integrate in vitro Caco-2 permeability measurements to make predictions of bioavailability that can be readily incorporated into risk assessment in general and
IVIVE in particular. While currently calibrated for Caco-2 transwell experiments, the general approach can be readily adapted to other types of in vitro or microphysiological models. Indeed, the high degree of chemical-to-chemical variability evident in the relationships between in vitro and in vivo permeability and between permeabilities across different gut segments points to the critical need for more accurate in vitro models for segment-specific gut absorption.

**References**


high-throughput heuristics for prioritizing human exposure to environmental chemicals. Environ Sci Technol 48, 12760-12767. doi:10.1021/acs.est.0c03538

Conflict of interest
The authors declare they have no conflicts of interest.

Data availability
All raw data and computational codes for modeling and data analysis are available in the supplementary materials1 and the GitHub repository2.

Acknowledgments
This work was supported, in part, by grants from the National Institute of Environmental Health Sciences (P42 ES027704, P30 ES029067) and the U.S. Environmental Protection Agency (RD 84003201). The views expressed in this manuscript do not reflect those of the funding agencies.