In silico PREDICTION OF ORAL BIOAVAILABILITY

Michael Lawless, John DiBella, Michael B. Bolger, Robert D. Clark, Eva Huehn, Marvin Waldman, Jinhua Zhang, and Viera Lukacova

Simulations Plus, Inc.
www.simulations-plus.com

ABSTRACT

Background. Oral bioavailability (F%) is an important pharmacokinetic property that can determine the fate of a compound in clinical trials. Predicting F% directly from the 2D structure of the molecule prior to first-in-human dosing is highly desirable.

Methods. A database of 62 drugs, primarily metabolized by cytochrome P450 (CYP) enzymes, including their F% and dose was created. Artificial neural network ensemble (ANNE) models based on 2D molecular structures were used to predict aqueous and biorelevant solubility, pK_a, logD, gastrointestinal permeability, fraction unbound in human plasma, and blood to plasma concentration ratio. A hierarchical set of models was used to determine CYP metabolism. First, classification models predicted whether each compound was a substrate for each of five major CYP isoforms (1A2, 2C9, 2C19, 2D6, and 3A4). Next, sites of metabolism were predicted for those compounds that were projected to be substrates. Finally, K_m and V_{max} predictions were made for each predicted site of metabolism. These predictions were used as inputs for physiologically based pharmacokinetic (PBPK) models implemented within GastroPlus™.

Results. All of the drugs were predicted to be well absorbed. All molecules were correctly predicted to be substrates of the CYPs associated with their major clearance pathways. Furthermore, these pathways had the highest predicted CYP intrinsic clearance in 42 of the 62 molecules. Overall, 68% of the molecules were predicted within 2-fold of their reported F%.

Conclusions. *In silico* tools predicting F% from 2D molecular structures can play an important role in anticipating challenges prior to initiating clinical development. Refining ANNE models based on the chemical space of interest is one strategy for improving prediction of F%.

INTRODUCTION

Oral bioavailability (F%) can determine the fate of a compound in clinical trials. Predicting F% directly from the 2D structure of the molecule prior to first-in-human dosing is highly desirable. We created a database of 62 drugs, primarily metabolized by CYP enzymes, that included the observed bioavailability (F%) and the recommended dose.[1] The reported F% values varied from 3% (fluphenazine) to 99% (diazepam, galantamine, glimepiride, indomethacin, and tamsulosin), with an average of 60%. See Figure 1 for representative data.

Drug	Dose [mg]	F%	MCP
irbesartan	150	70	2C9
vortioxetine	10	75	2D6
ibuprofen	400	85	2C9

Drug	Dose [mg]	F%	MCP
fluphenazine	5	3.1	2D6
verapamil	80	22	3A4
oxycodone	15	42	3A4

Figure 1. Examples of drugs in the data set along with their dose, F% and MCP.

METHODOLOGY

Artificial neural network ensemble (ANNE) models[2] based on 2D molecular structures were used to predict physicochemical properties and CYP pharmacokinetic parameters required as input into a PBPK model for each drug. These models were then used to predict the bioavailability solely from structure for each compound. The *in silico* predicted parameters included aqueous and biorelevant solubility, pK_a, logD, GI permeability, fraction unbound in human plasma, and blood to plasma concentration ratio. See Table 1.

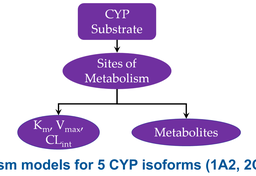


Figure 2. CYP metabolism models for 5 CYP isoforms (1A2, 2C9, 2C19, 2D6, and 3A4).

A hierarchical set of ANNE models was used to determine CYP metabolism. See Figure 2. The first model predicts whether a molecule is a substrate for each CYP isoform. These predictions include confidence estimates.[3] Next, sites of metabolism are predicted for compounds that are predicted as substrates. Finally, kinetic parameters are predicted and metabolites are depicted.

QSAR Model	Description
S+Sw	aqueous solubility
S+Sp	aqueous solubility at specified pH
S+FaSSGF	solubility in simulated fasted stated gastric fluid
S+FaSSIF	solubility in simulated fasted state intestinal fluid
S+FeSSIF	solubility in simulated fed state intestinal fluid
S+logD	logD at specified pH
S+pKa	pK_a (single or multiple)
S+Peff	effective human jejunal permeability
S+PrUnbnd	percent unbound to plasma proteins
S+RBP	blood-to-plasma concentration ratio
DiffCoef[4]	molecular diffusion coefficient in water
MET_XXX_Km	Kinetic Michaelis-Menten K_m constant (5 CYP isoforms)
MET_XXX_Vmax	Michaelis-Menten V_{max} constant (5 CYP isoforms)

Table 1. QSAR models used in PBPK simulations.

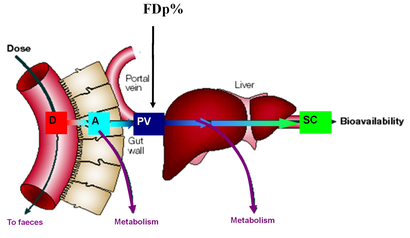
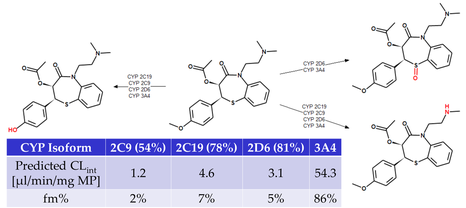


Figure 3[5]. Orally dosed drugs typically dissolve in the stomach and transit into the intestine, where they can be absorbed into the gut wall.

Fa% (fraction absorbed) is the fraction of dose that is absorbed into the apical membrane of the gut epithelium. CYP enzymes metabolize some compounds in the enterocytes. FDp% is the fraction/percent of dose that makes it to the portal vein. F% is the fraction/percent of dose that enters systemic circulation. Fa%, FDp%, and F% were predicted by our GastroPlus™ PBPK simulations.[6] A 35-year-old American male physiology was used for all PBPK simulations.

RESULTS



CYP Isoform	2C9 (54%)	2C19 (78%)	2D6 (81%)	3A4
Predicted CL_{int} [µl/min/mg MP]	1.2	4.6	3.1	54.3
fm%	2%	7%	5%	86%

Figure 4. Example of metabolite predictions for diltiazem.

Dilitiazem provides a good example of how this sort of analysis is applied. The drug is metabolized by several CYPs to yield several different metabolites in unequal amounts. The compound is predicted to be attacked by four of the major CYPs - 2C9, 2C19, 2D6 and 3A4 (Figure 4). Three sites of metabolism are predicted and the metabolites are displayed. The table contains the predicted intrinsic clearances and the fraction/percent metabolized (fm%) confidence estimates are shown in parenthesis.

The methoxy and dimethylamino groups are attacked by all four isoforms, whereas oxidation of the sulfur is only predicted for CYPs 2D6 and 3A4. Examining the corresponding CL_{ints}, however, indicates that CYP 3A4 will be responsible for the majority (86%) of the metabolism and that N-demethyldilitiazem will be the major (58%) metabolite.

Figure 5 shows the observed versus predicted F% for the 62 compounds in the database.

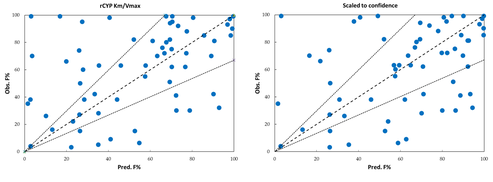


Figure 5. Observed vs. predicted F% for 62 compounds.

The dashed line is the line of unity. The dotted lines represent 1.5-fold errors. On the right-hand graph, V_{max} was scaled by the confidence estimate from the CYP substrate model (i.e., multiplied V_{max} x Confidence%/100), reducing severe underpredictions to avoid early rejection of good candidates.

Based on the modeling results, all of the drugs were predicted to be well absorbed. All molecules were correctly predicted to be substrates of the CYPs associated with their major clearance pathways. Furthermore, these pathways had the highest predicted CYP intrinsic clearance in 42 of the 62 molecules. Overall, 58% of the molecules were predicted within 1.5-fold of their reported F%.

For 46 compounds, the reported F% either spanned a range (e.g., 80-90%) or included a standard deviation (e.g., 80 ± 15%). See Figure 6. The area of the bubble is proportional to the expected measurement uncertainty. The F% of lovastatin is difficult to simulate due to opening and closing of the lactone ring. Tamsolusin has the highest fold error. NOTE: PBPK simulations using *in vitro* microsomal K_m and V_{max} values also resulted in large difference between reported and predicted F%. Propranolol was incorrectly predicted to be metabolized by 1A2; only including metabolism by 2D6 gives a correct F% prediction.

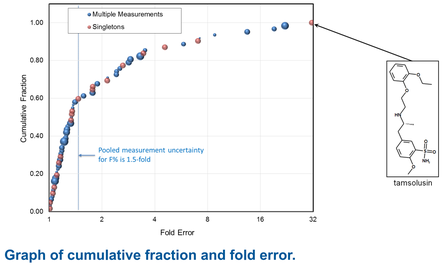


Figure 6. Graph of cumulative fraction and fold error.

DISCUSSION

In silico tools predicting F% from 2D molecular structures can play an important role in anticipating challenges prior to initiating clinical development. Refining ANNE models based on the chemical space of interest is one strategy for improving prediction of F%. It is also important to note that PBPK simulations using in vitro microsomal K_m and V_{max} values also resulted in large differences between reported and predicted F% for some compounds. This reflects uncertainties in the bioavailability measurements themselves as well simplifying assumptions in the model details, e.g., that the CYPs involved follow simple Michaelis-Menten kinetics across the entire concentration range, which is often not the case.

CONCLUSIONS

- A database of 62 drugs including oral bioavailability (F%) and dose was constructed
  - All compounds' reported major clearance pathways (MCP) were CYP-mediated[7]
- All molecules were predicted to be substrates of the CYP associated with their MCP
- In 42 of the 62 molecules, the CYP isoform with highest predicted intrinsic clearance (CL_{int}) was the same as the MCP
- Overall, 58% of the molecules were predicted within 1.5-fold of their reported F%
- Scaling V_{max} by the CYP substrate model's confidence estimate resulted in fewer underpredictions

REFERENCES

1. Thummel KE et al., In: Brunton LL, Chabner BA, Knollmann BC, editors. Goodman & Gilman's the pharmacological basis of therapeutics. 12th ed. New York: McGraw-Hill; 2011.
2. ADMET Predictor™ version 7.2, Simulations Plus, Inc., Lancaster, CA 95354 USA.
3. Clark RD et al., *J. Cheminform* **2014**, 6:34.
4. Hayduk W and Laudie H, *American Institute of Chemical Engineers J.* **1974**, 20:611.
5. Modified from van de Waterbeemd H and Gifford E. *ADMET In Silico Modelling: Towards Prediction Paradise?* Nat. Rev. Drug Disc. **2003**, 2:192:204.
6. GastroPlus™ version 9.0, Simulations Plus, Inc., Lancaster, CA 95354 USA.
7. Toshimoto K et al, *Drug Metabol. Disp*. Fast Forward. Published on August 14, 2014.

ACKNOWLEDGEMENTS

- The authors wish to acknowledge the assistance of Jill M. Harlach, Sandra Cohen, and Ted Grasela for their help in preparation of the poster.

Michael Lawless, John DiBella, Michael B. Bolger, Robert D. Clark, Eva Huehn, Marvin Waldman, Jinhua Zhang, and Viera Lukacova. In silico Prediction of Oral Bioavailability. Poster Presented at: American Society for Clinical Pharmacology and Therapeutics (ASCPT); March 8-12, 2016; San Diego, CA.

For additional information, please contact
Robert D. Clark
Simulations Plus
42505 10th Street West, Lancaster, CA 93534
(661) 723-7723 or bob@simulations-plus.com

This poster was previously presented at International Society for the Study of Xenobiotics