1521-009X/50/7/909–922$35.00
DRUG METABOLISM AND DISPOSITION

dx.doi.org/10.1124/dmd.121.000807
Drug Metab Dispos 50:909–922, July 2022

Selection of an Optimal In Vitro Model to Assess P-gp Inhibition: Comparison of Vesicular and Bidirectional Transcellular Transport Inhibition Assays[S]

Jocelyn Yabut,[1] Robert Houle,[1] Shubing Wang, Andy Liaw, Ravi Katwaru, Hannah Collier, Lucinda Hittle, and Xiaoyan Chu

Department of Pharmacokinetics, Pharmacodynamics and Drug Metabolism (J.Y., R.H., R.K., H.C., L.H., X.C.), and Department of Biometrics Research (S.W., A.L.), Merck & Co., Inc., Kenilworth, New Jersey

Received December 8, 2021; accepted April 4, 2022

ABSTRACT

The multidrug resistance protein 1 (MDR1) P-glycoprotein (P-gp) is a clinically important transporter. In vitro P-gp inhibition assays have been routinely conducted to predict the potential for clinical drug-drug interactions (DDIs) mediated by P-gp. However, high interlaboratory and intersystem variability of P-gp IC_{50} data limits accurate prediction of DDIs using static models and decision criteria recommended by regulatory agencies. In this study, we calibrated two in vitro P-gp inhibition models: vesicular uptake of N-methyl-quinidine (NMQ) in MDR1 vesicles and bidirectional transport (BDT) of digoxin in Lilly Laboratories Cell Porcine Kidney 1 cells overexpressing MDR1 (LLC-MDR1) using a total of 48 P-gp inhibitor and noninhibitor drugs and digoxin DDI data from 70 clinical studies. Refined thresholds were derived using receiver operating characteristic analysis, and their predictive performance was compared with the decision frameworks proposed by regulatory agencies and selected reference. Furthermore, the impact of various IC_{50} calculation methods and nonspecific binding of drugs on DDI prediction was evaluated. Our studies suggest that the concentration of inhibitor based on highest approved dose dissolved in 250 ml divided by $IC_{50}(I_2/IC_{50})$ is sufficient to predict P-gp related intestinal DDIs. IC_{50} obtained from vesicular inhibition assay with a refined threshold of $I_2/IC_{50} \geq 25.9$ provides comparable predictive power over those measured by net secretory flux and efflux ratio in LLC-MDR1 cells. We therefore recommend vesicular P-gp inhibition as our preferred method given its simplicity, lower variability, higher assay throughput, and more direct estimation of in vitro kinetic parameters, rather than BDT assay.

SIGNIFICANCE STATEMENT

This study has conducted comprehensive calibration of two in vitro P-gp inhibition models: uptake in MDR1 vesicles and bidirectional transport in LLC-MDR1 cell monolayers to predict DDIs. This study suggests that IC_{50}s obtained from vesicular inhibition with a refined threshold of $I_2/IC_{50} \geq 25.9$ provide comparable predictive power over those in LLC-MDR1 cells. Therefore, vesicular P-gp inhibition is recommended as the preferred method given its simplicity, lower variability, higher assay throughput, and more direct estimation of in vitro kinetic parameters.

Introduction

Multidrug resistance protein 1 (MDR1) P-glycoprotein (P-gp) is a clinically important transporter (Giacomini et al., 2010; Lee et al., 2010). Inhibition of P-gp can cause drug-drug interactions (DDIs), in which inhibition of intestinal P-gp appears to have the most significant impact (Fenner et al., 2009; Giacomini et al., 2010; Zhang et al., 2018). Thus, the U.S. Food and Drug Administration (FDA) and the European Medicines Agency (EMA) have requested evaluating the potential of a new molecular entity (NME) to inhibit P-gp in vitro and recommended a decision framework to determine the need for conducting clinical DDI studies with P-gp probe substrates, such as digoxin and dabigatran etexilate (DE) (EMA, 2012).

In vitro P-gp inhibition studies are routinely conducted in the pharmaceutical industry to evaluate the potential of NMEs as in vivo inhibitors of P-gp based on the recommendations from regulatory agencies. To assess their predictability for digoxin DDIs, a P-gp IC_{50} working group measured in vitro P-gp IC_{50} values for 15 compounds in 23 laboratories using their own assays and protocols. Substantial interlaboratory variability for IC_{50} values was reported (Bentz et al., 2013). A receiver operating characteristic (ROC) analysis was conducted by this group and refined cut-off values were proposed, which accounted for interlaboratory variability for IC_{50} values (Ellens et al.,

This work received no external funding. All studies conducted in this manuscript were sponsored by Merck Sharp & Dohme Corp., a subsidiary of Merck & Co., Inc., Kenilworth, NJ.

No author has an actual or perceived conflict of interest with the contents of this article.

[1]J.Y. and R.H. contributed equally to this work.

dx.doi.org/10.1124/dmd.121.000807.

[S] This article has supplemental material available at dmd.aspetjournals.org.

ABBREVIATIONS: AUC, area under the curve; AUC_{ROC}, area under the ROC curve; BCS, Biopharmaceutical Classification System; BDT, bidirectional transport; DDI, drug-drug interaction; DE, dabigatran etexilate; EMA, European Medicines Agency; ER, efflux ratio; FDA, United States Food and Drug Administration; FN, false negative; FP, false positive; fu[I_1], maximum concentration of inhibitor at steady state, fraction unbound in plasma; [I_{1u}], unbound [I_1]; [I_2], concentration of inhibitor in the gastrointestinal tract based on highest approved dose dissolved by 250 ml; LLCPK1-MDR1, Lilly Laboratories Cell Porcine Kidney 1 cells overexpressing MDR1; MDR1, multidrug resistance protein 1; NME, new molecular entity; NSF, net secretory flux; P_{app}, apparent permeability; P-gp, P-glycoprotein; PK, pharmacokinetic; ROC, receiver operating characteristic; TN, true negative; TP, true positive; UDF, unidirectional flux.

2013). Currently, P-gp IC_{50} variability is still a major concern that precludes accurate DDI prediction. As such, EMA has recommended assessing P-gp inhibition using two separate in vitro systems. A systemic calibration of in vitro assays in each individual laboratory may help to address this issue before the standardized model and assay protocol are established and employed. For instance, Cook et al. (2010), Sugimoto et al. (2011), and Poirier et al. (2014) calibrated their P-gp inhibition assays in Caco-2 cells (human colon adenocarcinoma cells expressing endogenous P-gp) and LLC-MDR1 cells (Lilly Laboratories Cell Porcine Kidney 1 cells overexpressing MDR1) using 26 to 68 clinical digoxin DDI data and defined their cut-off criteria accordingly.

In vitro P-gp inhibition is most frequently evaluated in P-gp transfected cells and Caco-2 cells. These polarized cells form a tight monolayer and therefore can be used to determine the inhibitory effect of a test compound on bidirectional transport (BDT) of a P-gp probe substrate, e.g., digoxin, from basolateral to apical (B to A) and from apical to basolateral (A to B) (Brouwer et al., 2013). However, this assay requires culturing cells for multiple days to form monolayers. Digoxin transcellular flux is not only mediated by P-gp and passive diffusion, but also by endogenous uptake transporter(s) (Taub et al., 2011; Lee et al., 2014). Varying expression of P-gp and endogenous transporters in different cell lines under different assay conditions may be one of the primary contributing factors to high interlaboratory variability. Furthermore, kinetic analysis of intrinsic inhibitory potency of NMEs on transcellular flux of digoxin by P-gp is complex. Thus, apparent IC_{50} values obtained from conventional analysis may not represent true P-gp inhibitory potency (Zamek-Gliszczynski et al., 2013; Jani and Krajcsi, 2014; Volpe et al., 2014). Alternatively, a vesicular uptake assay is a simpler non cell-based assay to study P-gp inhibition. The inhibitory effect of a test compound on ATP-dependent uptake of P-gp probe substrate, e.g., N-methyl-quinidine (NMQ), can be measured in MDR1 vesicles (Heredi-Szabo et al., 2013). Unlike the BDT assay, membrane vesicles can be stored in large quantitate to ensure consistent transporter expression/activity. As inside-out vesicles have direct access to P-gp binding sites, inhibition kinetics follows enzymatic principles. However, vesicular P-gp inhibition can only use the probe substrates with low permeability, e.g., NMQ. This can be a concern when extrapolating inhibition data to digoxin or other P-gp substrates, as P-gp has multiple binding sites (Lee et al., 2010; Sziraki et al., 2011). Currently, calibration of P-gp vesicular inhibition assay, their predictive performance, and interlaboratory variability is still limited (Ellens et al., 2013; Heredi-Szabo et al., 2013; Fekete et al., 2015).

In this study, we systemically evaluated the predictive performance of two in vitro P-gp inhibition models: vesicular uptake in MDR1 vesicles and BDT in LLC-MDR1 cells using a total of 48 drugs and digoxin DDI data from 70 clinical studies. Refined cut-off values using ROC analysis were derived for respective in vitro models and compared with the decision frameworks proposed by FDA, EMA, and Ellens et al. (2013). Furthermore, the impact of various IC_{50} calculation methods and nonspecific binding of inhibitor drugs on DDI prediction was evaluated.

Materials and Methods

Chemicals and Reagents

Chemicals. [^{3}H]digoxin (30–40 Ci/mmol) was purchased from Perkin Elmer (Boston, MA). [^{3}H] NMQ (L-000543643-002R001, 73 Ci/mmol) was synthesized by the Labeled Compound Synthesis Department, Merck & Co., Inc., Kenilworth, NJ. Sixty compounds evaluated in P-gp inhibition assays (48 compounds for training set and 12 compounds for test set) were obtained from Sigma (St. Louis, MO), Selleck Chemicals (Pittsburgh, PA), or Cayman Chemical (Ann Arbor, MI). All other reagents were commercially obtained at the highest analytical purity grade.

Cells and Membrane Vesicles. LLC-PK1 cells and LLC-PK1 cells stably expressing human MDR1 P-gp (LLC-MDR1 cells) were obtained from BD Gentest (Woburn, MA). LLC-MDR1 and LLC-PK1 cells obtained from Netherlands Cancer Institute (Amsterdam, Netherlands) were also used to measure in vitro IC_{50} values for several compounds in the training and test sets. Based on our internal validation, IC_{50} values measured using these two orthogonal cell lines show good correlation (data not shown). The cells were cultured in medium 199 supplemented with 10% fetal bovine serum, 2 mM L-glutamine, 50 U/ml penicillin, and 50 μg/ml streptomycin. All cells were maintained at 37°C in an atmosphere of 95% air, 5% CO_2, and 90% relative humidity. Membrane vesicle (lot EUD8G26 and IKATG03) isolated from baculovirus infected *Spodoptera frugiperda* (*Sf*9) cells containing MDR1 P-gp were obtained from Thermo Fisher Scientific (Waltham, MA). Similar time- and ATP-dependent uptake of [^{3}H] NMQ (0.1 μM) were observed in these two lots of vesicles (data not shown). The sideness of the vesicles was not measured. As nitrogen cavitation method was used to prepare membrane vesicles, we assumed that the vesicles consist of the mixture of equal portion of inside-out and right-side out vesicles (Saito et al., 2009).

In Vitro P-gp Inhibition Assays

BDT Inhibition Assay in LLC-MDR1 and LLC-PK1 Cells. The effects of test compounds on human MDR1 P-gp-mediated efflux transport of digoxin were evaluated using the BDT studies in the LLC-MDR1 and LLC-PK1 cell lines as previously described (Chan et al., 2019). Briefly, cells were cultured in 96-well multiwell insert plates (Millicell-96, Millipore, Billerica, MA) at 85,000 cells/well and cultured for 4 days before the study. The compound was tested at seven concentrations in LLC-MDR1 cells as indicated. Cyclosporin A (CsA) (10 μM) was used as a positive control inhibitor. The test compounds or positive control inhibitor at the concentrations indicated were added into both apical (A) and basolateral (B) sides of cell monolayers. Transport buffer was Hanks' balanced salt solution with 10 mM (N-[2-Hydroxyethyl] piperazine-N'-[2-ethanesulfonic acid (HEPES), pH 7.4. Transport of digoxin was measured in both absorptive and secretory directions. For absorptive (A to B) transport, the donor dosing solution was added to the apical compartment, and for secretory (B to A) transport, donor dosing solution was added to the basolateral compartment. Receiver solution was prepared by adding aliquots of the stock solution of test compound or positive control inhibitor to transport buffer with a final organic solvent concentration of $\leq$ 1%. Donor dosing solution was prepared by diluting aliquots of radiolabeled and nonradiolabeled [^{3}H]digoxin (final concentration 0.1 μM), and, if applicable, aliquots of test compound or positive control inhibitor stock solutions were added into transport buffer at designated concentrations with a final organic solvent concentration of $\leq$ 1%. BDT of digoxin without inhibitor was tested in both LLC-MDR1 and LLC-PK1 cells to confirm P-gp-mediated digoxin efflux transport and a sufficient signal-to-noise ratio (efflux ratios [ER, B to A/A to B of apparent passive permeability (P_{app})] in LLC-MDR1 cells/ER in LLC-PK1 cells $\geq$ 3). Prior to the transport experiment, cells were washed three times with transport buffer. Donor dosing solution (150 μl) was added to either the apical or basolateral compartment, with receiver solution (150 μl) added to the opposite compartment. At 90 minutes, (50 μl) samples were taken from both sides and scintillant (Ultima Gold, Perkin Elmer, Boston, MA) was added. Radioactivity was determined by liquid scintillation counting in a 2450 MicroBeta2 counter (Perkin Elmer, Waltham, MA). Dextran Texas Red or Lucifer Yellow was used as the markers to test the monolayer integrity. At the end of the incubation, if Dextran Texas Red or Lucifer Yellow in the receiver well was > 6% or > 2% of the total concentration, respectively, data were excluded due to poor monolayer integrity. The experiments were performed in triplicate.

Vesicular Transport Inhibition Assay in MDR1 P-gp Containing Membrane Vesicles. The inhibitory effect of test compounds on ATP-dependent [^{3}H]NMQ (0.1 μM) uptake was measured in membrane vesicles containing human MDR1 P-gp. The positive control inhibitor (CsA 10 μM) was tested in each assay to confirm the functionality of MDR1 P-gp. Briefly, 19 μl of [^{3}H]NMQ, dissolved in transport buffer (0.25 M sucrose, 10 mM $MgCl_2$, 10 mM Tris-HCl buffer, pH 7.4), were added to 10 μl of MDR1 P-gp containing vesicles (2.5 mg/ml) in a 96-well glass coated plate (Analytical Sales & Services, Flanders, NJ). Then, 1 μl of various concentrations of test compounds or CsA (10 μM) were added to each corresponding well. The incubation plate was preincubated in a water bath for 3 minutes at 37°C. Uptake was initiated by the addition of 20 μl ATP or AMP containing solution (final concentration of 5 mM ATP or AMP, 10 mM creatine phosphate, and 100 μg/ml creatine phosphokinase in

transport buffer), followed by incubation at 37°C for 5 minutes. Uptake was stopped by the addition of 200 μl ice-cold stop buffer (0.25 M sucrose, 0.1 M NaCl, 10 mM Tris-HCl buffer, pH 7.4), followed by transfer of the reaction mixture to a prewetted 96-well glass fiber filter plate (1.0 μm) (Millipore, Billerica, MA) and application of vacuum. Filters containing the membrane vesicles were washed with 200 μl ice-cold stop buffer five times. The filter plate was dried at room temperature, and 100 μl scintillation fluid (OptiPhase HiSafe3; PerkinElmer, Boston, MA) was added. Radioactivity was determined by liquid scintillation counting. The experiments were performed in triplicate.

Quantification of Test Compounds Via Liquid Chromatography Tandem Mass Spectrometry Analysis

All analyses were performed on a Sciex 4500 triple quadrople mass spectrometer (Toronto, ON, Canada) equipped with electrospray source using multiple reaction monitoring and controlled by Analyst 1.6.2 software. The sample was loaded onto an Acquity UPLC C18 HSS T3, 1.8 μM, 2.1 × 30 mm, 1.8 micron column (Waters, Milford, MA) by means of a Thermo Scientific LX-2 System (Leap Autosampler with Dionex Ultimate 3000 RS Pumps) autosampler controlled by Aria 1.7 software. Chromatography was performed using water, 0.1% formic acid as mobile phase A and acetonitrile, 0.1% formic acid as mobile phase B. The fast gradient profile was 0 to 15 seconds 5% (v/v) B at 0.8 ml/min; 30 seconds ramp to 95% (v/v) B at 1.0 ml/min; 30 seconds to 5% (v/v) B at 0.8 ml/min. Polarity was selected for optimum sensitivity and detection by tandem mass spectrometry was based on precursor ion transitions to the strongest intensity product ions. Samples, standards, and controls were processed with labetalol as the internal standards. Typical standard curve range was from 7.82 to 2000 nM. Liquid chromatography tandem mass spectrometry (LC/MS/MS) parameters were shown in Supplemental Table 1. Data processing was done using IndigoBio Ascent (Indigo BioAutomation, Carmel, IN).

Clinical DDI Data with Digoxin and Other P-gp Probe Substrates

Clinical data for 70 DDI studies using digoxin as a probe drug (Table 1, training set) and 18 and 6 studies using DE and fexofenadine as probe drugs, respectively (Table 2) were collected from the University of Washington DDI database (https://www.druginteractionsolutions.org). A diverse compound set were categorized according to the Provisional Biopharmaceutical Classification System (BCS) obtained from the literature (Wu and Benet, 2005; Benet, 2013; Papich and Martinez, 2015), University of Washington DDI database, and Pharmapendium (https://www.pharmapendium.com). The magnitude of clinical DDIs expressed as AUCR, the ratio of area under the curve (AUC) or $C_{max}R$, the ratio of maximum concentration in plasma (C_{max}) [pharmacokinetic (PK) ratios] with and without inhibitor drugs were collated. A PK ratio of a probe drug ≥ 1.25 was considered a positive clinical DDI, whereas a PK ratio < 1.25 was defined as a negative outcome. I_1, the mean steady-state total (free and bound) C_{max}, and the fraction unbound in plasma (f_u) for test compounds were collected from the University of Washington DDI database. Extrapolated I_1 values were used, assuming the linear PK, if the values at indicated inhibitor doses were not reported. I_{1u} is the unbound I_1; f_u of 0.01 was used in DDI risk assessment as the worst-case scenario if reported $f_u < 0.01$. I_2, the concentration of drug in the gastrointestinal tract at the given inhibitor dose dissolved in 250 ml was calculated. In addition, a test set of 12 compounds was assembled to evaluate performance of DDI prediction, and related clinical DDI data of the compounds in test sets with orally administered digoxin was collected as indicated above and shown in Supplemental Table 2.

Data Analysis

IC_{50} Calculation

BDT inhibition assay. IC_{50} values for inhibiting digoxin BDT in LLC-MDR1 cells were obtained by fitting the data to eq. 1 by nonlinear regression analysis using Graphpad prism (San Diego, CA).

$$\%\text{Control} = 100/(1 + I^s/IC_{50}{}^s) \quad (1)$$

where I is the inhibitor concentration (μM), and s is the Hill slope.
Percent control was calculated according to eq. 2:

$$\%\text{Control} = (T_I/T_o) \times 100 \quad (2)$$

where T_I represents net transport of digoxin measured in the presence of various concentrations of the inhibitor, and T_o represents the net transport of digoxin in the absence of the inhibitor.
Three methods were commonly used to calculate the net transport of digoxin in LLC-MDR1 cells: 1) net secretory flux (NSF), 2) efflux ratio (ER), and 3) unidirectional flux (UDF) (Balimane et al., 2008; Cook et al., 2010; O'Connor et al., 2015). The NSF evaluates net digoxin transport activity in both absorptive and secretory directions. The ER describes the ratio of P_{app} in secretory (P_{app} B to A) over the absorptive direction (P_{app} A to B). The UDF (B to A) describes transport activity of digoxin in the secretory direction (basolateral to apical: B to A).

NSF in LLC-MDR1 cells was calculated according to eq. 3:

$$\text{NSF} = (\%\text{Transport B to A}) - (\%\text{Transport A to B}). \quad (3)$$

%Transport was calculated by dividing the amount of digoxin measured in the receiver compartment by the sum of digoxin measured in both receiver and donor compartments.

ER in LLC-MDR1 cells was calculated via eq. 4:

$$\text{ER} = P_{app}(\text{B to A})/P_{app}(\text{A to B}). \quad (4)$$

P_{app} is calculated by eq. 5:

$$P_{app} = (\text{Volume of receiver chamber})/(\text{A x } C_o) \times \Delta\text{Conc}/\Delta t. \quad (5)$$

The volume of receiver chamber is 0.15ml; transwell membrane area (A) is $0.11cm^2$; C_o is the sum of the probe substrate concentration (digoxin) measured in the donor plus receiver compartments at 1.5 hours; Δ in concentration is the concentration in the receiver compartment at 1.5 hours; and Δ in Time is the incubation time (1.5 × 60 × 60 = 5400s). P_{app} is expressed as 10^{-6} cm/s.
UDF (B to A) in LLC-MDR1 cells was calculated by eq. 6:

$$\text{UDF (B to A)} = (P_{app}BA_i - P_{app}BA_p)/(P_{app}BA_o - P_{app}BA_p) \quad (6)$$

where P_{app} BA_i represents P_{app} receiver B to A with inhibitor; $P_{app}BA_o$, B to A receiver without inhibitor, P_{app} BA_p, B to A receiver with positive control inhibitor (CsA, 10 μM).
Considering potential nonspecific binding of inhibitor compounds to assay plates and cells, a separate set of IC_{50} analysis was performed by correcting IC_{50} values obtained from the nominal inhibitor concentrations based on the recovery of inhibitor drugs measured in the incubation.

The recovery of the inhibitor drugs (R_i) was measured in well via LC/MS/MS analysis under the BDT assay conditions as described above. R_i was calculated by dividing the sum of the inhibitor concentration measured in the donor (C_{donor}) and receiver ($C_{receiver}$) compartments at the end of the incubation divided by C_{dose}, the concentration of test compound in dosing solution at time zero. The concentrations of the inhibitor drugs remained in the cells were not measured.

$$R_i = \left[(C_{donor} + C_{receiver})/C_{dose}\right]. \quad (7)$$

As the recovery of the test compound was calculated only based on the concentration of the test compound in apical and basolateral chambers, lower recovery suggested possible nonspecific binding of the test compound to cell monolayers and/or the assay plates. If R_i was < 0.7, in vitro IC_{50} values measured based on nominal inhibitor concentrations were corrected based on eqs. 8–10, respectively. If R_i was ≥ 0.7, the IC_{50} values were not corrected.

TABLE 1

Clinical digoxin DDI data used in ROC analysis
All clinical data in this table were used for ROC analysis as training set.

Inhibitor	BCS[a]	Dose (mg)	I_1 (μM)[b]	f_u[c]	I_2 (μM)	AUCR	$C_{max}R$	Clinical DDIs Y/N[d]
Alogliptin	III	25	0.44	0.80	295	1.00	0.94	N
Amiodarone	II	400	3.5	0.0002	2479	1.63	1.72	Y
		600	5.3	0.0002	3719	1.69	1.75	Y
		800	7.0	0.0002	4960	1.68	1.84	Y
Apixiban	III	20	0.94	0.13	174	0.90	0.92	N
Asunaprevir	II	200	0.49	0.03	1069	1.30	1.09	Y
Atorvastatin	II	10	0.013	0.02	72	1.00	1.10	N
		80	0.12	0.02	573	1.10	1.20	N
Azilsartan	IV	80	9.24	0.01	563	1.03	0.94	N
Bosentan	II	500	5.90	0.02	3626	0.87	0.94	N
Canagliflozin	IV	300	7.60	0.0125	2597	1.20	1.36	Y
Captopril	III	12.5	0.50	0.75	230	1.39	1.59	Y
Carvedilol (Male[e])	II	6.5	0.13	0.05	62	1.24	1.00	N
Carvedilol (Female[e])	II	6.5	0.13	0.05	62	1.56	1.38	Y
Clarithromycin	II	500	3.20	0.54	2674	1.46	1.75	Y
Clopidogrel	II	75	0.012	0.02	932	1.02	1.10	N
Daclatasvir	II	60	2.34	0.01	325	1.27	1.65	Y
Diltiazem	I	60	0.17	0.22	579	1.51	1.37	Y
		90	0.49	0.22	868	1.40	1.38	Y
		180	0.20	0.22	1737	1.50	1.37	Y
Dronedarone	II	400	0.19	0.02	2874	2.57	1.75	Y
Elagolix	III	200	1.23	0.2	1267	1.26	1.71	Y
Elbasvir	IV	50	0.14	0.001	227	1.11	1.47	Y
Eliglustat	I	100	0.18	0.38	989	1.37	1.68	Y
		150	0.27	0.38	1479	1.41	1.64	Y
Etravirine	IV	200	2.20	0.01	1838	1.18	1.19	N
Felodipine	II	2.5	0.007	0.004	26	1.49	1.35	Y
		5	0.013	0.004	52	1.11	1.38	Y
		10	0.034	0.004	104	1.16	1.39	Y
Fidaxomicin	IV	200	0.005	0.022	756	1.12	1.18	N
Flibanserin	II	100	1.07	0.02	1025	1.93	1.46	Y
Isradipine	II, IV	5	0.03	0.03	54	1.11	1.26	Y
Itraconazole	II	200	0.90	0.002	1134	1.68	1.34	Y
Ivacaftor	II, IV	150	13.90	0.02	1529	1.32	1.23	Y
Maraviroc	III	300	0.52	0.24	2336	1.02	1.09	N
Mibefradil	II	50	0.92	0.004	404	1.08	1.22	N
		100	1.85	0.004	807	1.07	1.25	Y
		150	2.77	0.004	1211	1.31	1.41	Y
Mirabegron	III	100	1.00	0.29	1009	1.27	1.29	Y
Nelfinavir	II	1250	7.04	0.015	8806	1.35	1.34	Y
Nicardipine	I	20	0.11	0.0125	167	1.07	NA	N
		30	0.17	0.0125	250	0.96	0.90	N
Nifedipine	I, II	5	0.12	0.04	58	1.21	1.01	N
		10	0.23	0.04	115	1.23	1.06	N
		20	0.46	0.04	231	1.18	1.08	N
Nitrendipine	II	10	0.01	0.02	111	1.09	1.22	N
		20	0.02	0.02	222	1.15	1.57	Y
Paroxetine	I	30	0.18	0.05	364	0.85	0.90	N
Quinidine	I	200	2.00	0.13	2466	NA	1.42	Y
		250	2.50	0.13	3082	NA	2.18	Y
		600	6.00	0.13	7398	2.65	NA	Y
Ranolazine	II	400	3.36	0.38	3742	1.39	2.30	Y
		750	6.68	0.38	7017	1.88	1.68	Y
		1000	8.40	0.38	9356	1.60	1.46	Y
Ritonavir	II	200	5.00	0.015	1110	1.22	1.26	Y
		400	10.00	0.015	2219	1.39	1.25	Y
Rolapitant	II	180	1.93	0.002	1439	1.30	1.71	Y
Rosuvastatin	II	40	0.02	0.12	332	1.04	1.01	N
Sertraline	I, II	200	0.41	0.015	2612	1.10	1.05	N
Telaprevir	II	750	5.24	0.325	4413	1.85	1.50	Y
Telmisartan	II	120	1.12	0.005	933	1.22	1.50	Y
Ticagrelor	IV	400	6.99	0.02	3062	1.28	1.75	Y
Troglitazone	II	400	3.00	0.01	3624	1.04	1.05	N
Valspodar	IV	200	1.49	0.025	659	3.05	2.44	Y
		400	1.56	0.025	1317	1.74	1.72	Y
Vandetanib	II	300	3.32	0.06	2524	1.23	1.29	Y
Velpatasvir	IV	100	0.70	0.005	453	1.27	1.87	Y
Vemurafenib	IV	960	125.30	0.01	7838	1.91	1.42	Y

TABLE 1 *continued*

Inhibitor	BCS[a]	Dose (mg)	I_1 (μM)[b]	f_u[c]	I_2 (μM)	AUCR	$C_{max}R$	Clinical DDIs Y/N[d]
Verapamil	I	80	0.37	0.1	704	1.50	1.44	Y
		120	0.55	0.1	1056	NA	1.61	Y

f_u, the fraction unbound in plasma; NA, Not available.
[a]BCS classification of inhibitor drugs was obtained from the University of Washington DDI database (https://www.druginteractionsolutions.org), https://www.pharmapendium.com, and the literature (Wu and Benet, 2005; Benet, 2013; Papich and Martinez, 2015).
[b]Clinical DDI data, I_1, and f_u values were collected from the University of Washington DDI database (https://www.druginteractionsolutions.org); I_1 reported from clinical DDI studies was used. Otherwise, I_1 at same or similar inhibitor dose regimen was collected. In the cases that I_1 at indicated inhibitor doses was not reported, extrapolated I_1 was used, assuming the linear PK.
[c]Either AUCR and/or $C_{max}R$ (PK ratios) greater or equal to 1.25 were considered a positive clinical DDI. If f_u measured < 0.01, f_u value of 0.01 was used for DDI prediction.
[d]Yes (Y) indicates AUCR or $C_{max}R \geq 1.25$; No (N) indicates AUCR and $C_{max}R < 1.25$.
[e]Male and female cohort tested.

$$IC_{50}(NSF_c) = IC_{50}(NSF) \times R_i \quad (8)$$

$$IC_{50}(ER_c) = IC_{50}(ER) \times R_i \quad (9)$$

$$IC_{50}(UDF_c) = IC_{50}(UDF) \times R_i \quad (10)$$

where IC_{50}(NSF), IC_{50}(ER), and IC_{50}(UDF) represent IC_{50} values estimated by calculating digoxin net transport using NSF, ER, and UDF, respectively. $IC_{50}(NSF_c)$, $IC_{50}(ER_c)$, and $IC_{50}(UDF_c)$ represent corrected IC_{50} values based on the recovery of test compounds.

Vesicular Inhibition Assay

MDR1 P-gp-mediated NMQ transport was estimated based on ATP-dependent [^{3}H] NMQ uptake in MDR1 vesicles calculated by subtracting uptake of [^{3}H] NMQ in presence of AMP from uptake in presence of ATP. Percent control in vesicular inhibition assay is calculated according to eq. 11:

$$\%Control = (V_I/V_o) \times 100 \quad (11)$$

where V_I represents ATP-dependent uptake rate of NMQ measured in the presence of various concentrations of the test compound, and V_o represents ATP-dependent uptake rate in the absence of the test compound. As [^{3}H] NMQ exhibited minimal active uptake in control vesicles and in MDR1 vesicles in the presence of ATP and positive control inhibitor (10 μM CsA) (Supplemental Fig. 1), no control vesicles or positive control inhibitor were used to correct P-gp-mediated NMQ transport rate.

The IC_{50} for inhibition of MDR1 P-gp-mediated NMQ vesicular transport, IC_{50}(V), was obtained by fitting the data to eq. 1.

Inhibitor drugs may have nonspecific binding to the assay plates and vesicles. As our vesicular inhibition assay was conducted in glass-coated plates and the surface area as well as the volume of vesicles are small (Nervi et al., 2010) at the given vesicle amount (25 μg) in this assay, nonspecific binding of inhibitor drugs may have negligible

TABLE 2

Selected clinical P-gp DDI data using dabigatran etexilate and fexofenadine as probe drugs
Only clinical DDI data with inhibitor drug IC_{50} data shown in Table 1 were collected.

Inhibitor	BCS[a]	Inhibitor Dose (mg)	I_1 (μM)[b]	I_2 (μM)	f_u[b]	Clinical Probe[c]	Probe Drug Dose (mg)	AUCR[b]	$C_{max}R$[b]	Clinical DDIs (Y/N)[d]
Amiodarone	II	600	2.43	3720	0.002	DE	150	1.6	1.5	Y
Atorvastatin	II	80	0.12	573	0.026	DE	150	0.82	0.8	N
Clarithromycin	III	500	0.002	2674	0.5	DE	150	1.6	1.49	Y
Clarithromycin	III	500	3.2	2674	0.5	DE	0.375	4.22	4.58	Y
Clopidogrel	II	75	0.008	932	0.02	DE	150	0.92	0.95	N
Clopidogrel	II	300	0.055	3729	0.02	DE	150	1.36	1.68	Y
Clopidogrel	II	600	0.055	7458	0.02	DE	150	1.32	1.43	Y
Cobicistat	II	150	0.0015	773	0.053	DE	150	2.4	2.33	Y
Dronedarone	II	400	0.2	2874	0.02	DE	150	2.36	2.25	Y
Itraconazole	II	200	0.75	1134	0.002	DE	0.375	7.4	6.42	Y
Quinidine	I	200	0.56	2466	0.13	DE	150	1.53	1.56	Y
Rifampin	II	600	18.96	2916	0.4	DE	0.375	2.38	1.78	Y
Ritonavir	IV	100	2	555	0.31	DE	150	1.11	1.13	N
Ticagrelor	IV	90	0.899	689	0.016	DE	150	1.56	1.46	Y
Ticagrelor	IV	180	2.76	1378	0.016	DE	150	1.73	1.95	Y
Verapamil	II	120	0.13	1056	0.225	DE	150	1.39, 2.08[e]	1.33, 2.29[e]	Y
Verapamil	II	240	0.26	2112	0.225	DE	150	1.71	1.91	Y
Alogliptin	III	100	2.058	1179	0.694	fexofenadine	80	1.26	1.07	Y
Itraconazole	II	200	0.483	1134	0.002	fexofenadine	180	2.29	2.69	Y
Paroxetine	I	20	0.02	4000	0.05	fexofenadine	60	1.38	1.33	Y
Quinidine	I	200	NA	2466	0.13	fexofenadine	25	2.14	2.39	Y
Sertraline	I	50	NA	653	0.01	fexofenadine	50	0.84	0.86	N
Verapamil	II	240	NA	2112	0.225	fexofenadine	120	1.46, 2.5[e]	1.3, 2.9[e]	Y

f_u, the fraction unbound in plasma.
[a]BCS classification of inhibitor drugs was obtained from the University of Washington DDI database (https://www.druginteractionsolutions.org), https://www.pharmapendium.com, and the literature (Wu and Benet, 2005; Benet, 2013; Papich and Martinez, 2015).
[b]Clinical DDI data, I_1, and f_u values were collected from the University of Washington DDI database (https://www.druginteractionsolutions.org); I_1 reported from clinical DDI studies was used. Otherwise, I_1 at same or similar inhibitor dose regimen was collected. In the cases that I_1 at indicated inhibitor doses was not reported, extrapolated I_1 was used, assuming the linear PK. If f_u measured < 0.01, f_u value of 0.01 was used for DDI prediction.
[c]DE, the prodrug of dabigatran was administered for dabigatran DDI studies.
[d]Either AUCR and/or $C_{max}R$ (PK ratios) greater or equal to 1.25 were considered a positive clinical DDI; Yes (Y) indicates AUCR or $C_{max}R \geq 1.25$; No (N) indicates AUCR and $C_{max}R < 1.25$.
[e]Multiple clinical DDI data reported.

impact on inhibitor concentrations in the incubation (nominal extravesicular concentration). This was confirmed by our pilot studies for several highly bound inhibitor drugs after the measurement of the binding of inhibitor drugs to vesicles using ultracentrifugation method (unpublished observations). Therefore, in this study, IC_{50} estimated based on nominal inhibitor concentrations was used for vesicular inhibition assay without the correction for nonspecific binding.

Statistical Methods and the Criteria to Define Optimized Cut-Off Values

A binary classification analysis was conducted on clinical digoxin DDI studies for the compounds in the training set (Table 1) to derive the optimal cut-off values to predict digoxin DDIs based on three static models: 1) I_2/IC_{50} (FDA, 2020); 2) I_1/IC_{50} or I_2/IC_{50} (Agarwal et al., 2013); and 3) I_{1u}/IC_{50}, or I_2/IC_{50} (EMA, 2012).

IC_{50} values measured in the BDT and vesicular inhibition assays and calculated using various methods [IC_{50}(NSF), $IC_{50}(NSF_c)$, IC_{50}(ER), $IC_{50}(ER_c)$, IC_{50}(UDF), $IC_{50}(UDF_c)$, and IC_{50}(V)] were used in this analysis, and the predictive performances were compared.

The possible outcomes were: 1) true positive (TP), in vitro data predicts a positive digoxin DDI and the prediction is in agreement with a positive clinical DDI (AUCR $\geq$ 1.25 or $C_{max}R \geq 1.25$); 2) false negative (FN), in vitro data predict a negative digoxin DDI, but the prediction is not in agreement with positive clinical data; 3) true negative (TN), in vitro data predict a negative DDI and the prediction is in agreement with a negative clinical DDI (AUCR < 1.25 and $C_{max}R < 1.25$); and 4) false positive (FP), in vitro data predict a positive DDI, but it is not observed in clinic.

Performance metrics used in the analyses are defined and calculated as below:

$$\text{Sensitivity} = \text{True positive rate} = \text{TP}/\text{P} \quad (12)$$

$$\text{Specificity} = \text{True negative rate} = \text{TN}/\text{N} \quad (13)$$

$$\text{Overall accuracy} = (\text{TP} + \text{TN})/\text{S} \quad (14)$$

$$\text{Average accuracy} = (\text{Sensitivity} + \text{Specificity})/2 \quad (15)$$

$$\text{False negative rate} = \text{FN}/\text{P} \quad (16)$$

$$\text{False positive rate} = 1 - \text{Specificity} \quad (17)$$

S is the total number of digoxin DDI studies, P the number of positive clinical DDI studies (AUCR $\geq$ 1.25 or $C_{max}R \geq 1.25$), and N is the number of negative studies (AUCR < 1.25 and $C_{max}R < 1.25$).

The receiver operating characteristic (ROC) curve, which is a plot with the sensitivity versus false positive rate (1-specificity) at various cutoffs, is used to check and visualize the overall performance of a binary classifier. The AUC of a ROC curve (AUC_{ROC}) is one of the most important evaluation metrics. An $AUC_{ROC} = 1$ indicates a perfect separation of the two classes (positive or negative), and an $AUC_{ROC} = 0.5$ means that the classifier is the same as randomly splitting the outcomes. We first apply ROC analysis on the one-dimensional classifiers: the static models of I_2/IC_{50}(model 1). The ROC curve is generalized to the two-dimensional classifiers for static model I_1/IC_{50} or I_2/IC_{50}(model 2) and I_{1u}/IC_{50} or I_2/IC_{50}(model 3), respectively, by plotting the sensitivity versus the sorted false positive rate after pooling all the possible two-dimensional cutoffs.

In the analysis, the AUC_{ROC}s (Hanley and McNeil, 1982; Robin et al., 2011) and their 95% confidence intervals (Hanley and McNeil, 1982) were used to evaluate and compare the overall classification power of proposed classifiers/tests. The classifier with the best AUC_{ROC} was selected. The computation was conducted using open-source programming language **R** (http://www.R-project.org).

A cutoff is one or a set of values that defines a positive or negative outcome in a binary classifier. Therefore, the cutoff also defines the performance metrics. ROC analysis uses these metrics to determine the optimal cutoff. For a chosen classifier, the following method was used to find the optimal cutoffs (C):

$$C = argmax_{c \in \{\theta : Sensitivity(\theta) \geq 0.75\}} \text{Specificity}\ (c) \quad (18)$$

where "*argmax*" is an operation that finds the argument that gives the maximum value from a target function. *Sensitivity* (c) and *Specificity* (c) are defined for a given cutoff c. This method maximizes *Specificity* by searching in the cutoff space, where a *Sensitivity* of $\geq$ 0.75 is guaranteed.

The predictive performance using derived cut-offs based on ROC analysis was compared with the following static models and recommended cut-off values by regulatory agencies and Ellens et al. (2013): 1) $I_2/IC_{50} \geq 10$ (FDA, 2020); 2) $I_2/IC_{50} \geq 45$ (Ellens et al., 2013); 3) $I_1/IC_{50} \geq 0.1$ or $I_2/IC_{50} \geq 10$ (Agarwal et al., 2013); 4) $I_1/IC_{50} \geq 0.03$ or $I_2/IC_{50} \geq 45$ (Ellens et al., 2013); and 5) $I_{1u}/IC_{50} \geq 0.02$ or $I_2/IC_{50} \geq 10$ (EMA, 2012). The cut-offs recommended by FDA and EMA were empirical values, whereas the cut-off proposed by Ellens et al. (2013) was derived based on ROC analysis of P-gp IC_{50} data for 15 compounds generated by 23 laboratories using four in vitro systems: Caco-2 cells, LLCPK1-MDR1, MDCKII-MDR1, and MDR1 vesicles. P-gp probe substrates were digoxin for polarized cell-lines and NMQ or vinblastine for MDR1 vesicles (Ellens et al., 2013). After optimizing the cut-off values for different models using the data from the training set, the performance of selected models and optimized cut-offs were further verified on a test set with 12 compounds not included in the training set.

Results

Comparison of In Vitro P-gp IC_{50} Values Measured in the BDT and Vesicular Inhibition Assays

In vitro IC_{50} values of 48 compounds in the training set measured in BDT and vesicular inhibition assays are summarized in Table 3, and IC_{50} plots of all test compounds measured in BDT assay using NSF, ER, and MDR1 vesicular inhibition assay are shown in Supplemental Fig. 2-1 and 2–2. For the BDT assay, IC_{50} values were calculated using NSF, ER, and UDF, with and without correction for nonspecific binding. Obtained IC_{50} data [IC_{50}(NSF), $IC_{50}(NSF_c)$, IC_{50}(ER), $IC_{50}(ER_c)$, IC_{50}(UDF), and $IC_{50}(UDF_c)$] were compared with IC_{50} values from vesicular inhibition assay [IC_{50}(V)], respectively (Fig. 1). In cases where the IC_{50} values could not be determined due to the lack of more than 50% inhibition (not an inhibitor at highest tested concentration or the solubility limit), the highest inhibitor concentrations tested were used as surrogate IC_{50} values for the purpose of comparison across different assays and methods.

The variability of vesicular versus BDT inhibition assay was evaluated. Supplemental Table 3 compared the variability the IC_{50} values for verapamil and quinidine conducted in three independent studies in both MDR1 vesicles and LLC-MDR1 cells [IC_{50}(V) versus IC_{50}(NSF)]. Based on this limited dataset, IC_{50} values measured in MDR1 vesicles were less variable (lower CV%) than in BDT assay. This is also supported by the data collected in the literature when the same assay systems and probe substrates were used. We have also compared IC_{50}(V) values generated in this study with those reported in the literature using MDR1 vesicles. The results were summarized in Supplemental Table 4. The IC_{50} values reported in literature using MDR1 vesicles are still limited (25 out of 48 compounds have IC_{50} data reported in the literature). Nevertheless, 6 out of 25 compounds showed > 10-fold difference (10.8–22-fold) on IC_{50}(V)s compared with those reported values. Despite that, such interlaboratory variability is lower than reported by

TABLE 3

In vitro IC_{50} values of 48 compounds measured in P-gp bidirectional and vesicular transport inhibition assays
All data shown in this table were used as training set for ROC analysis.

Compounds	IC_{50}(NSF)	IC_{50}(NSF_c)	IC_{50}(ER)	IC_{50}(ER_c)	IC_{50}(UDF)	IC_{50}(UDF_c)	IC_{50}(V)
	(μM)						
Alogliptin	>100	>100	>100	>100	~207.0 ± 91.0	~207.0 ± 91.0	>1000
Amiodarone	21.9 ± 2.6	0.6 ± 0.1	8.3 ± 2.6	0.2 ± 0.1	10.6 ± 0.6	0.3 ± 0.02	14.8 ± 0.7
Apixaban	>150	>150	>150	>150	>150	>150	>150
Asunaprevir	21.2 ± 2.0	21.2 ± 2.2	17.3 ± 2.7	17.3 ± 2.7	23.7 ± 3.0	23.7 ± 3.0	3.2 ± 0.2
Atorvastatin	>100	>100	>100	>100	>100	>100	17.0 ± 1.7
Azilsartan	13.2 ± 2.7	2.6 ± 0.5	4.3 ± 0.3	0.8 ± 0.1	20.9 ± 2.5	4.2 ± 0.5	8.2 ± 0.7
Bosentan	>15	>15	>15	>15	>15	>15	>15
Canagliflozin	59.0 ± 10.2	2.9 ± 0.5	39.9 ± 9.0	2.0 ± 0.5	>100	> 5	>100
Captopril	>1000	>67	>1000	>67	>1000	>67	>1000
Carvedilol	19.1 ± 1.9	11.0 ± 0.9	1.6 ± 0.1	0.5 ± 0.04	16.4 ± 2.4	5.5 ± 0.8	4.1 ± 0.3
Clarithromycin	~145.7 ± 26.6	~145.7 ± 26.6	>100	>100	>100	>100	10.9 ± 2.4
Clopidogrel	>100	>36	132.8 ± 50.0	48.4 ± 18.3	>100	>36	157.4 ± 15.1
Daclatasvir	>100	>5	35.5 ± 6.9	1.8 ± 0.3	>100	>5	1.6 ± 0.1
Diltiazem	83.2 ± 14.8	83.2 ± 14.8	62.7 ± 11.6	62.7 ± 11.6	23.7 ± 7.2	23.7 ± 7.2	30.5 ± 2.8
Dronedarone	82.0 ± 29.0	3.1 ± 1.1	7.1 ± 1.6	0.3 ± 0.1	>300	>11	4.9 ± 0.5
Elagolix	~168.7 ± 51.4	~168.7 ± 51.4	100.5 ± 31.0	100.5 ± 31.0	>150	>150	24.8 ± 0.8
Elbasvir	>0.5	>0.5	>0.5	>0.5	>0.5	>0.5	0.3 ± 0.02
Eliglustat	>100	>100	43.7 ± 12.4	43.7 ± 12.4	>100	>100	65.4 ± 2.1
Etravirine	>10	>2.1	>10	>2.1	>10	>2.1	>30
Felodipine	11.4 ± 0.9	3.6 ± 0.2	3.8 ± 0.6	0.8 ± 0.01	>50	>10.5	81.8 ± 5.1
Fidaxomicin	>100	>100	>100	>100	>100	>100	0.4 ± 0.05
Flibanserin	51.4 ± 5.7	51.4 ± 5.7	8.8 ± 0.7	8.8 ± 0.7	~141.3 ± 76.3	~141.3 ± 76.3	>120
Isradipine	29.5 ± 1.7	29.5 ± 1.7	6.1 ± 0.3	6.1 ± 0.3	62.8 ± 16.6	62.8 ± 16.6	53.7 ± 2.4
Itraconazole	6.9 ± 0.7	1.2 ± 0.1	1.1 ± 0.3	0.2 ± 0.04	2.4 ± 0.2	0.4 ± 0.04	0.34 ± 0.13
Ivacaftor	1.8 ± 0.6	0.3 ± 0.1	2.1 ± 1.0	0.4 ± 0.2	0.6 ± 0.1	0.1 ± 0.04	> 1.0
Maraviroc	>1000	>1000	>1000	>1000	>1000	>1000	162.6 ± 8.5
Mibefradil	7.8 ± 1.8	1.9 ± 0.4	5.2 ± 0.4	1.3 ± 0.1	14.0 ± 5.2	3.4 ± 1.3	10.0 ± 1.6
Mirabegron	>300	>300	>300	>300	>300	>300	148.9 ± 21.8
Nelfinavir	14.8 ± 3.0	2.7 ± 0.5	1.3 ± 0.2	0.3 ± 0.03	7.3 ± 1.3	1.4 ± 0.2	20.3 ± 1.6
Nicardipine	5.3 ± 0.6	1.5 ± 0.2	0.8 ± 0.1	0.2 ± 0.02	4.1 ± 1.6	1.2 ± 0.5	6.4 ± 1.7
Nifedipine	83.8 ± 9.8	14.2 ± 2.6	18.3 ± 0.8	3.1 ± 0.1	58.9 ± 19.0	10.0 ± 3.2	115.6 ± 6.4
Nitrendipine	20.0 ± 4.1	20.0 ± 4.1	3.1 ± 0.1	3.1 ± 0.1	>100	> 100	76.0 ± 1.4
Paroxetine	>100	> 21.5	9.3 ± 0.9	2.0 ± 0.3	>100	> 21.5	122.2 ± 15.3
Quinidine	56.0 ± 9.5	56.0 ± 9.5	12.5 ± 2.2	12.5 ± 2.2	59.0 ± 12.8	59.0 ± 12.8	14.5 ± 1.9
Ranolazine	74.5 ± 9.6	74.5 ± 9.6	13.2 ± 0.7	13.2 ± 0.7	215.6 ± 62.0	215.6 ± 62.0	64.1 ± 4.1
Ritonavir	18.9 ± 1.0	18.9 ± 1.0	14.7 ± 0.6	14.7 ± 0.6	> 15	>15	0.3 ± 0.03
Rolapitant	22.2 ± 6.3	5.6 ± 0.4	6.7 ± 0.5	1.6 ± 0.1	34.9 ± 15.6	8.4 ± 3.7	>30
Rosuvastatin	>300	>300	>300	>300	>300	>300	>300
Sertraline	30.6 ± 12.8	7.7 ± 3.2	6.3 ± 1.0	1.6 ± 0.2	6.5 ± 0.8	1.7 ± 0.2	39.4 ± 2.8
Telaprevir	>200	>200	>200	>200	>200	>200	4.1 ± 0.5
Telmisartan	20.3 ± 2.8	20.3 ± 2.8	1.6 ± 0.2	1.6 ± 0.2	25.9 ± 8.9	25.9 ± 8.9	0.8 ± 0.1
Ticagrelor	~11.2 ± 3.1	~0.8 ± 0.3	3.7 ± 0.01	0.9 ± 0.002	~12.0 ± 6.1	~3.0 ± 1.5	>30
Troglitazone	17.1 ± 1.1	3.1 ± 0.2	9.3 ± 1.7	1.7 ± 0.3	13.5 ± 1.1	2.4 ± 0.2	19.4 ± 1.3
Valspodar	0.7 ± 0.1	0.1 ± 0.003	0.4 ± 0.1	0.04 ± 0.01	1.2 ± 0.3	0.05 ± 0.004	0.2 ± 0.01
Vandetanib	>10	>0.2	5.9 ± 0.9	0.1 ± 0.02	>10	>0.2	96.1 ± 9.1
Velpatasvir	>50	>2.5	>50	>2.5	>50	>2.5	4.9 ± 0.7
Vemurafenib	60.2 ± 25.0	8.8 ± 6.0	34.0 ± 15.0	5.3 ± 2.0	1.7 ± 0.9	0.3 ± 0.2	>30
Verapamil	39.6 ± 3.9	39.6 ± 3.9	3.2 ± 0.2	3.2 ± 0.2	52.9 ± 11.1	52.9 ± 11.1	2.8 ± 0.3

IC_{50}(NSF), IC_{50}(ER), and IC_{50}(UDF) represent IC_{50} values measured in bidirectional transport (BDT) assay estimated by calculating digoxin net transport using net secretory flux, efflux ratio, and unidirectional flux, respectively. IC_{50}(NSF_c), IC_{50}(ER_c), and IC_{50}(UDF_c) represent corrected IC_{50} values based on the recovery of inhibitor drugs measured in BDT assays. IC_{50}(V) represent IC_{50} values measured for inhibition of ATP-dependent NMQ vesicular transport. ~ (tilde operator) indicates approximation of IC_{50} extrapolated beyond maximum concentration tested. Data were reported as mean ± S.D. for triplicate measurement.

others in BDT assays using different P-gp transfected cells and Caco-2 cells (Bentz et al., 2013). Further studies are needed to confirm this observation.

To compare IC_{50}(NSF) with IC_{50}(V) (Fig. 1A-1), the IC_{50}(NSF) values for 14 out of 48 compounds were 5-fold higher than the IC_{50}(V), in which 10 compounds had at least 10-fold higher IC_{50}(NSF) than IC_{50}(V), ranging from 10.2-fold for velpatasvir to 224-fold for fidaxomicin. In contrast, IC_{50}(V) values for only 3 compounds were 5-fold higher than IC_{50}(NSF), ranging from 7.2-fold for felodipine to 10-fold for alogliptin. The difference for alogliptin was attributed to the difference of the highest concentrations tested, as no inhibition was observed in both assays at the highest concentrations tested. As shown in Fig. 1C-1, IC_{50}(ER) for 12 compounds were more than 5-fold lower than IC_{50}(V) (ranging from 6- to 24-fold), whereas IC_{50}(ER) values of 9 compounds were at least 5-fold higher than IC_{50}(V) [ranging from 5- to 224-fold (fidaxomicin)]. IC_{50}(UDF) values of 15 compounds were at least 5-fold higher than IC_{50}(V), whereas only 3 compounds had 5-fold lower IC_{50}(UDF) than IC_{50}(V) (Fig. 1 E-1). Overall, there is a poor correlation between IC_{50}(NSF) and IC_{50}(V) (Fig. 1A-2), IC_{50}(ER), and IC_{50}(V) (Fig. 1C-2), as well as IC_{50}(UDF) and IC_{50}(V) (Fig. 1E-2). The correlation was improved, when alogliptin and maraviroc were excluded from the analysis (Fig. 1, A-3, C-3, and E-3).

In the BDT assays, the IC_{50} values obtained after correction for nonspecific binding of inhibitor compounds [IC_{50}(NSF_c), IC_{50}(ER_c), and IC_{50}(UDF_c)] were compared with IC_{50}(V), respectively, and the results are shown in Fig. 1, B-1, D-1, and F-1 and Table 3. As described in the *Method* section, nonspecific binding of inhibitor drugs was not corrected for IC_{50}(V). The IC_{50} values corrected for nonspecific binding were

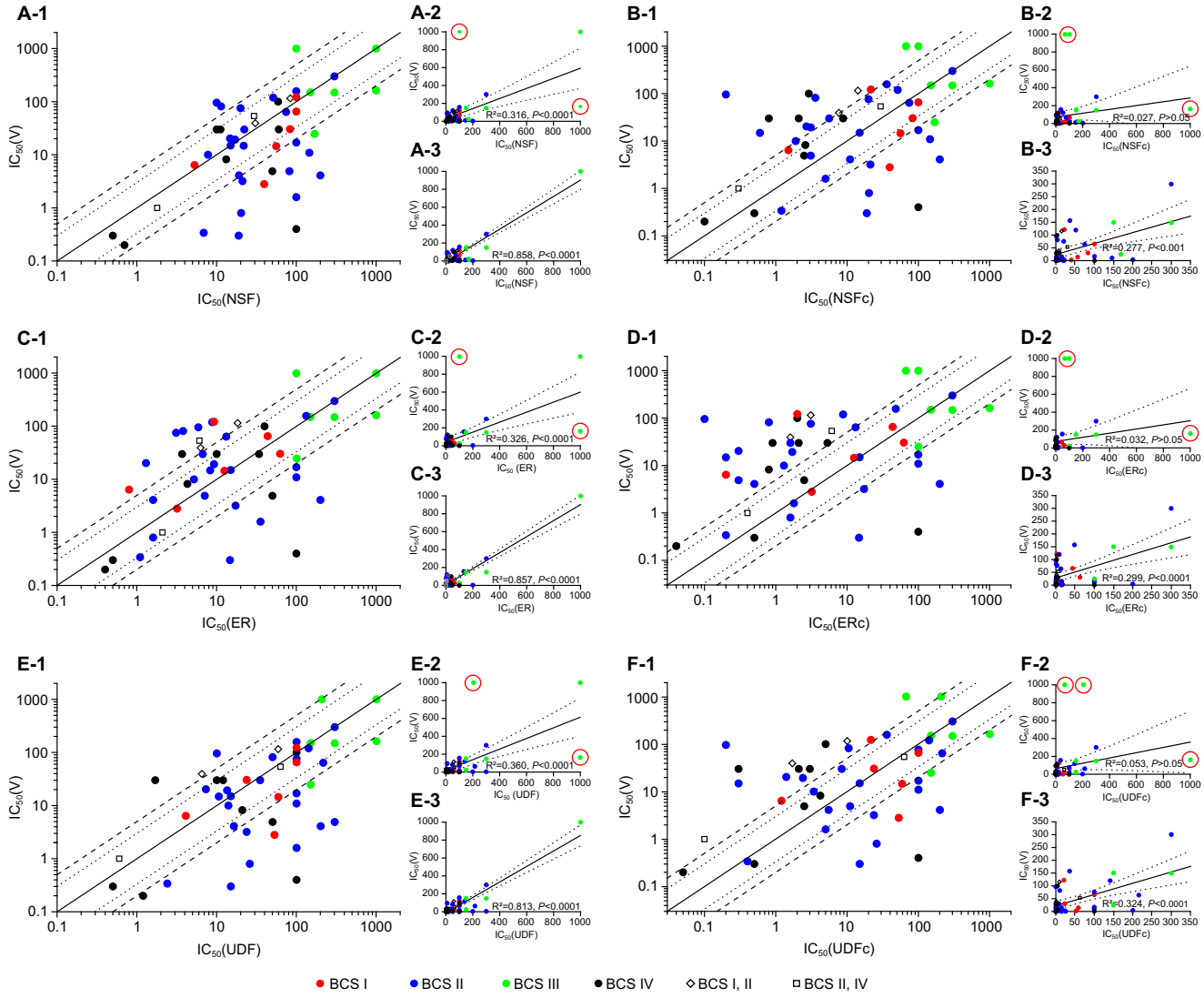

Fig. 1. Comparison of in vitro IC_{50} values of 48 compounds measured in bidirectional and vesicular transport inhibition assays. Panel A: IC_{50}(NSF) versus IC_{50}(V); Panel B: $IC_{50}(NSF_c)$ versus IC_{50}(V); C: IC_{50}(ER) versus IC_{50}(V); D: $IC_{50}(ER_c)$ versus IC_{50}(V); E: IC_{50}(UDF) versus IC_{50}(V); F: $IC_{50}(UDF_c)$ versus IC_{50}(V). The symbols in red, blue, green, and black circles represent the compounds classified as BCS I, II, III, and IV compounds, respectively; whereas the symbols in opened diamonds and squares represent the compounds that are classified as either BCS I or II, and BCS II or IV, respectively. In Panels A-1, B-1, C-1, D-1, E-1, and F-1, the solid, dotted, and dashed lines represent the line of unity and 3-fold and 5-fold differences, respectively. Panels A-2, B-2, C-2, D-2, E-2, and F-2 are correlation plots of respective data. Panels A-3, B-3, C-3, D-3, E-3, and F-3 are the correlation plots excluding the compounds in red circles [alogliptin and maraviroc in all figures (A-2 to F-2)], and captopril in B-2, D-2, and F-2. In all correlation plots, solid and dotted lines represent the regression line and its 95% confidence interval, respectively. All IC_{50} values (mean ± S.D.) were shown in Table 3.

generally lower for the compounds that had high nonspecific binding to the assay plates and/or cells as evidenced by low recovery in the incubation. The stability of test compound during incubation was confirmed (data not shown). $IC_{50}(NSF_c)$ for 15 compounds were 5-fold lower than IC_{50}(V). On the contrary, $IC_{50}(NSF_c)$ values for 10 compounds were 5-fold higher than IC_{50}(V), in which 6 compounds had 10-fold higher $IC_{50}(NSF_c)$ than IC_{50}(V) (Fig. 1B-1; Table 3). Compared with IC_{50}(ER), $IC_{50}(ER_c)$ values trended toward lower than IC_{50}(V): $IC_{50}(ER_c)$ of 23 compounds were more than 5-fold lower than IC_{50}(V), where 18 out of 23 compounds showed at least 10-fold lower $IC_{50}(ER_c)$ than IC_{50}(V). In contrast, only 7 compounds showed 5-fold higher $IC_{50}(ER_c)$ than IC_{50}(V). Similarly, $IC_{50}(UDF_c)$ for 15 compounds were more than 5-fold lower than IC_{50}(V), whereas $IC_{50}(UDF_c)$ of 10 compounds were more than 5-fold higher than IC_{50}(V). Overall, the correction for nonspecific binding did not improve, but rather reduced the correlation between the IC_{50} values measured by the BDT and vesicular inhibition assays. There was no correlation between $IC_{50}(NSF_c)$ and IC_{50}(V) (Fig. 1B-2; $R^2 = 0.027$, $P > 0.05$); $IC_{50}(ER_c)$ and IC_{50}(V) (Fig. 1D-2; $R^2 = 0.032$, $P > 0.05$), as well as $IC_{50}(UDF_c)$ and IC_{50}(V) (Fig. 1F-2, $R^2 = 0.053$, $P > 0.05$). Excluding alogliptin, maraviroc, and captopril resulted in an improved, but yet poor correlation (Fig. 1B-3; $R^2 = 0.277$, $P < 0.001$; Fig. D-3, $R^2 = 0.299$, $P < 0.001$; Fig. F-3, $R^2 = 0.324$, $P < 0.001$). Among the 48 compounds tested, the difference in observed IC_{50} values using the BDT and vesicular assays for

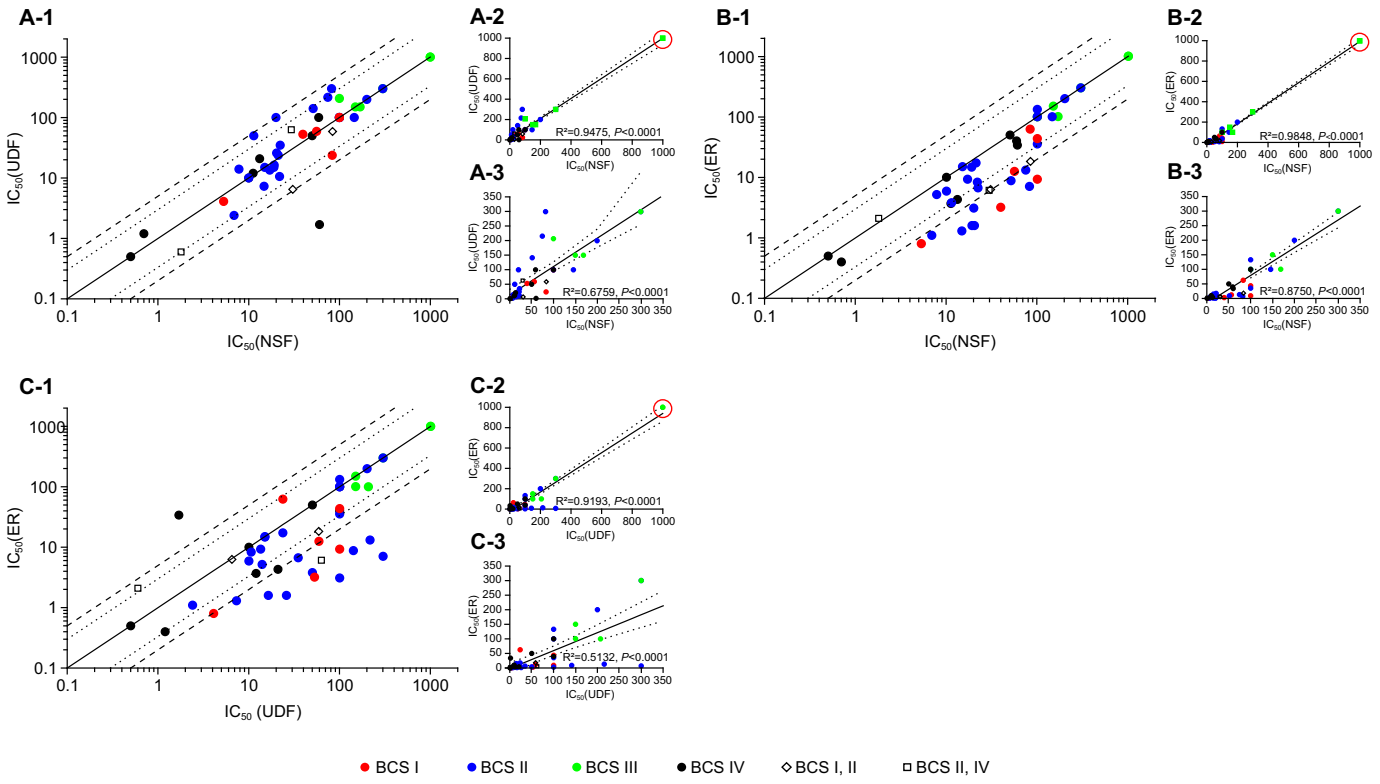

Fig. 2. Comparison of IC_{50} values measured in bidirectional transport assays using different calculation methods. Panel A: IC_{50}(NSF) versus IC_{50}(UDF); Panel B: IC_{50}(NSF) versus IC_{50}(ER); Panel C: IC_{50}(UDF) versus IC_{50}(ER). The symbols in red, blue, green, and black circles represent the compounds classified as BCS I, II, III, and IV compounds, respectively; whereas the symbols in opened diamonds and squares represent the compounds that are classified as either BCS I or II, and BCS II or IV, respectively. In Panels A-1, B-1, and C-1, the solid, dotted, and dashed lines represent the line of unity and 3-fold and 5-fold differences, respectively. Panels A-2, B-2, and C-2 are correlation plots of respective data. Panels A-3, B-3, and C-3 are the correlation plots excluding the compounds in red circles (captopril and maraviroc). In all correlation plots, solid and dotted lines represent the regression line and its 95% confidence interval, respectively. All IC_{50} values (mean ± S.D.) were shown in Table 3.

BCS I and III compounds appeared to be less than those for BCS II and IV compounds, despite the lack of a clear trend and a limited dataset.

Comparison of IC_{50} Values in the BDT Assays Using Different IC_{50} Calculation Methods

Fig. 2 showed IC_{50} values measured in the BDT assays calculated by NSF, ER, and UDF. IC_{50}(NSF) of all test compounds were comparable to those of IC_{50}(UDF) with the difference less than 3- to 5-fold, except for one outlier, vemurafenib (BCS IV compound), whose IC_{50}(NSF) was 35-fold higher than IC_{50}(UDF) (Fig. 2A-1; Table 3). IC_{50}(ER) exhibited a trend of lower than IC_{50}(NSF) (Fig. 2B-1; Table 3): IC_{50}(ER) values of 11 compounds were more than 5-fold lower than IC_{50}(NSF), in which 6 compounds were more than 10-fold lower than those of IC_{50}(NSF) (Fig. 2B-1; Table 3). Likewise, IC_{50}(ER) values also trended toward lower than IC_{50}(UDF), except vemurafenib's IC_{50}(ER), which was 20-fold higher than IC_{50}(UDF). IC_{50}(UDF) of 12 compounds were more than 5-fold higher than IC_{50}(ER), where 10 out of 12 compounds had more than 10-fold higher IC_{50}(UDF) (Fig. 2C-1). In general, there is a good correlation when IC_{50}(NSF) versus IC_{50}(UDF), IC_{50}(NSF) versus IC_{50}(ER), and IC_{50}(UDF) versus IC_{50}(ER) were compared. Excluding captopril and maraviroc, two compounds with IC_{50} values [IC_{50}(NSF), IC_{50}(ER), IC_{50}(UDF)] greater than 1000 μM, there was still a good correlation between IC_{50}(NSF) and IC_{50}(ER) (R^2= 0.875, $P < 0.0001$), followed by IC_{50}(NSF) versus

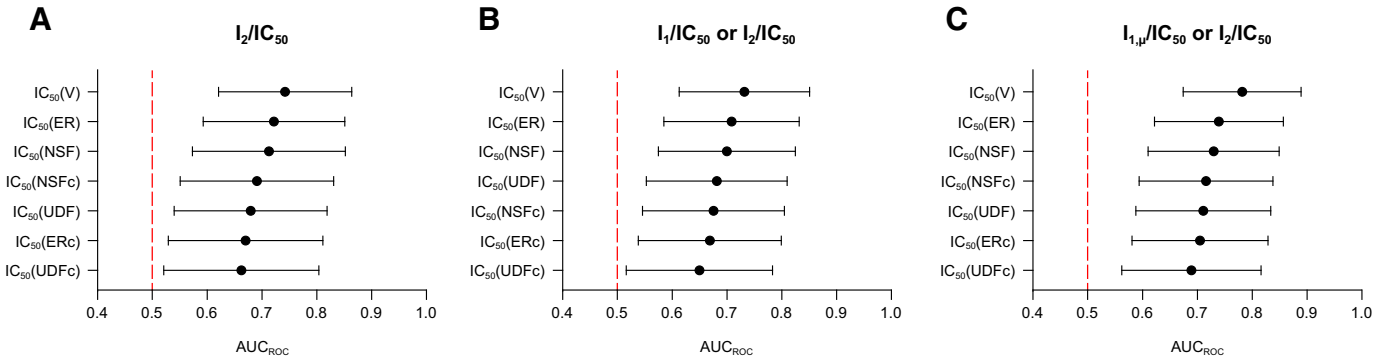

Fig. 3. The AUC_{ROC} values obtained from ROC analysis of different static models using various IC_{50} measurement. Estimated AUC_{ROC} values and their 95% confidence intervals of various static models were shown in panel A (I_2/IC_{50}), panel B (I_1/IC_{50}, or I_2/IC_{50}), and panel C (I_{1u}/IC_{50}, or I_2/IC_{50}). In vitro IC_{50} values [IC_{50}(V), IC_{50}(ER), IC_{50}(NSF), IC_{50}(UDF), IC_{50}(ER_c), IC_{50}(NSF_c), IC_{50}(UDF_c)] were measured as described in the *Methods and Materials* section.

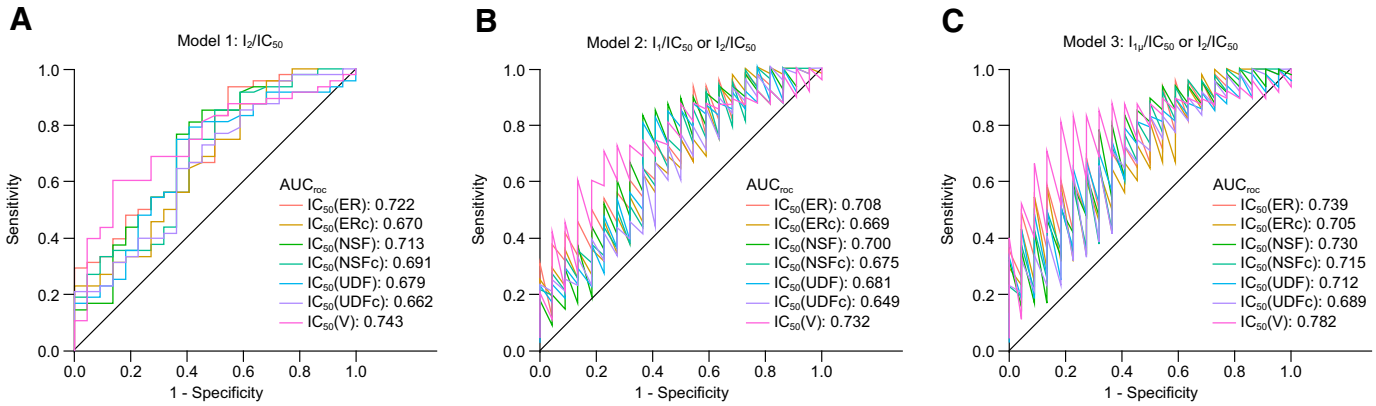

Fig. 4. ROC curves of various static models using different IC_{50} estimation methods. ROC curves with different IC_{50} measurement and calculation methods were shown in panel A (model 1: I_2/IC_{50}), panel B (model 2: I_1/IC_{50}, or I_2/IC_{50}), and panel C (model 3: I_{1u}/IC_{50}, or I_2/IC_{50}).

IC_{50}(UDF) (R^2= 0.676, $P < 0.001$), and IC_{50}(UDF) versus IC_{50}(ER) (R^2= 0.513, $P < 0.0001$) (Fig. 2, A-3, B-3, and C-3).

Clinical Digoxin DDI Data for ROC Analysis Training and Test Sets

Table 1 summarized 70 clinical digoxin DDI studies, I_1, I_2, f_u, AUCR, and $C_{max}R$ for 48 compounds with measured in vitro P-gp IC_{50} values (Table 3). These data were used as the training set for ROC analysis. Only clinical DDI data with orally administered digoxin were collected. For certain compounds, multiple clinical DDI datasets at different perpetrator dose regimen and study design were included. In this dataset, 48 clinical DDIs were positive (AUCR and/or $C_{max}R \geq 1.25$) and 22 DDI data were negative (AUCR and $C_{max}R < 1.25$). For 48 compounds tested, 6, 23, 6, and 9 compounds were classified as BCS Class I, II, III, and IV, respectively. Due to inconsistent information from BCS classification database and literature, 2 compounds were classified as BCS Class I or II, and another 2 compounds were classified as BCS Class II or IV. We also conducted in vitro P-gp inhibition studies for an additional 12 compounds as the test set. In vitro IC_{50} values obtained [IC_{50}(V), IC_{50}(ER), IC_{50}(NSF)] were shown in Supplemental Table 5 and related clinical digoxin DDI data (8 positive and 4 negative) are summarized in Supplemental Table 2. These data were used as the test set to assess the predictive performance of the models and the cut-off values derived by ROC analysis.

Comparison of ROC Analysis Using Various Static P-gp DDI Prediction Models and IC_{50} Measurement

To select the optimal model and IC_{50} assay to predict digoxin DDIs, ROC analysis was conducted with three P-gp DDI prediction models: 1) I_2/IC_{50}; 2) I_1/IC_{50} or I_2/IC_{50}; and 3) I_{1u}/IC_{50}, or I_2/IC_{50} using IC_{50} values [IC_{50}(V), IC_{50}(NSF), IC_{50}(NSF_c), IC_{50}(ER), IC_{50}(ER_c), IC_{50}(UDF), and IC_{50}(UDF_c)] obtained from various assays and calculation methods. AUC_{ROC} values and respective ROC curves obtained from this analysis were shown in Fig. 3 and Fig. 4, respectively. Despite the lack of statistical significance (possibly due to the relatively small sample sizes), model 1 (I_2/IC_{50}) with IC_{50}(V) demonstrated a trend of highest AUC_{ROC}, followed by IC_{50}(ER), IC_{50}(NSF), and IC_{50}(UDF). Similar observations were also found in model 2 (I_1/IC_{50}, or I_2/IC_{50}), and model 3 (I_{1u}/IC_{50}, or I_2/IC_{50}) (Fig. 3). In all three models, IC_{50} measurement in the BDT assay after the correction of nonspecific binding of inhibitor drugs [IC_{50}(ER_c), IC_{50}(NSF_c), and IC_{50}(UDF_c)] showed a trend of lower AUC_{ROC} than respective IC_{50} measurement without the correction of nonspecific binding [IC_{50}(ER), IC_{50}(NSF), and IC_{50}(UDF)]. As shown in Fig. 4, the ROC curve was well defined for one dimensional binary classifier (model 1: I_2/IC_{50}), but not for two dimensional classifiers (models 2 and 3: I_1/IC_{50} or I_2/IC_{50}; I_{1u}/IC_{50} or I_2/IC_{50}), as their true positive rate and true negative rate were not monotonically related.

Based on ROC analysis, optimal discrimination thresholds that achieve the highest specificity constrained to a minimal sensitivity of 0.75 was obtained for these models, and their predictive performance were summarized in Table 4. Only IC_{50}(V), IC_{50}(ER), and IC_{50}(NSF) values were used in this analysis, as they had relatively higher AUC_{ROC} values. The predictive performance of each model with optimal thresholds derived from the ROC analysis was compared with respective models and the cut-off values recommended by regulatory agencies (Agarwal et al., 2013; FDA, 2020; EMA, 2012) and Ellens et al. (2013) (Table 4).

In model 1 (I_2/IC_{50}), the cut-off value for I_2/IC_{50} using IC_{50}(V) was 25.9 with the percentage of TP, TN, FP, and FN of 75%, 59.1%, 40.9%, and 25%, respectively, and the average and overall accuracy of 0.67 and 0.7 (Table 4-1). Using the same data set, the cut-off values recommended by FDA ($I_2/IC_{50} \geq 10$) resulted in lower FN% (16.7%) but higher FP% (54.5%), whereas the cut-off by Ellens et al. (2013) ($I_2/IC_{50} \geq 45$) showed higher FN% (31.3%). Using IC_{50}(ER), the obtained cut-off value of 13.7 yielded the percentage of TP, TN, FP, and FN of 83.3%, 50%, 50%, and 16.7%, respectively, and the average and overall accuracy of 0.667 and 0.729, respectively. These results were generally comparable to those using FDA cut-off value ($I_2/IC_{50} \geq 10$), whereas the cut-off value by Ellens et al. (2013) resulted in a higher FN% (35.4%), despite a relatively lower FP% (36.4%). Likewise, the cut-off value derived based on IC_{50}(NSF) ($I_2/IC_{50} \geq 9.3$) is comparable to the one from FDA ($I_2/IC_{50} \geq 10$) with similar accuracy, and the percentage of TP, TN, FP, and FN, whereas the cut-off value by Ellens et al. (2013) ($I_2/IC_{50} \geq 45$) resulted in lower accuracy and higher FN%.

In model 2 (I_1/IC_{50}, or I_2/IC_{50}), the cut-off values obtained with IC_{50}(V), IC_{50}(ER), and IC_{50}(NSF) was (0.032, 40), (0.081, 26.7), and (0.026, 10), respectively (Table 4-2). The accuracy with IC_{50}(V) and IC_{50}(NSF) was the same, whereas the accuracy of IC_{50}(ER) was lower with higher FP% (50%). Using the cut-off value of $I_1/IC_{50} \geq 0.1$, or $I_2/IC_{50} \geq 10$, all three IC_{50} dataset had similar accuracy with lower FN% for IC_{50}(V) and IC_{50}(ER). Using the cut-off value of $I_1/IC_{50} \geq 0.03$, or $I_2/IC_{50} \geq 45$ by Ellens et al. (2013), IC_{50}(V) had the same FP% and FN% as the cut-off derived from ROC analysis, but IC_{50}(ER) and IC_{50}(NSF) resulted in lower accuracy and higher FN%. In addition, we have further compared the cut-off values and predictive performance derived from our ROC analysis using IC_{50}(V) data with those reported by Ellens et al. (2013) using only MDR1 vesicular IC_{50} dataset for 15 compounds generated in five laboratories using either NMQ or vinblastine as in vitro probes [data shown in Supplemental Table 2 of Ellens

TABLE 4

Summary of the cut-off values obtained from ROC analysis based on static P-gp DDI prediction models using various P-gp IC_{50} methods and the comparison of predictive performance with other cut-off criteria

Table 4-1: Model 1 (I_2/IC_{50})

Model 1 (I_2/IC_{50})	IC_{50}(V)[a]			IC_{50}(ER)[a]			IC_{50}(NSF)[a]		
	ROC Analysis	FDA , 2020[b]	Ellens et al., 2013[c]	ROC Analysis	FDA , 2020[b]	Ellens et al., 2013[c]	ROC Analysis	FDA , 2020[b]	Ellens et al., 2013[c]
Cut-off values	25.9	10	45	13.7	10	45	9.3	10	45
TP % (sensitivity)	75 (36/48)	83.3 (40/48)	68.8 (33/48)	83.3 (40/48)	85.4 (41/48)	64.6 (31/48)	77.1 (37/48)	75 (36/48)	52.1 (25/48)
TN % (specificity)	59.1 (13/22)	45.5 (10/22)	72.7 (16/22)	50 (11/22)	45.5 (10/22)	63.6 (14/22)	63.6 (14/22)	63.6 (14/22)	72.7 (16/22)
FP %	40.9 (9/22)	54.5 (12/22)	27.3 (6/22)	50 (11/22)	54.5 (12/22)	36.4 (8/22)	36.4 (8/22)	36.4 (8/22)	27.3 (6/22)
FN %	25 (12/48)	16.7 (8/48)	31.3 (15/48)	16.7 (8/48)	14.6 (7/48)	35.4 (17/48)	22.9 (11/48)	25 (12/48)	47.9 (23/48)
Average Accuracy	0.67	0.644	0.707	0.667	0.654	0.641	0.704	0.693	0.624
Overall Accuracy	0.7	0.714	0.7	0.729	0.729	0.643	0.729	0.714	0.586

Table 4-2: Model 2 (I_1/IC_{50} or I_2/IC_{50})

Model 2 (I_1/IC_{50} or I_2/IC_{50})	IC_{50}(V)[a]			IC_{50}(ER)[a]			IC_{50}(NSF)[a]		
	ROC Analysis	Agarwal et al., 2013[d]	Ellens et al., 2013[c]	ROC Analysis	Agarwal et al., 2013[d]	Ellens et al., 2013[c]	ROC Analysis	Agarwal et al., 2013[d]	Ellens et al., 2013[c]
Cut-off values	(0.032, 40)	(0.1, 10)	(0.03, 45)	(0.081, 26.7)	(0.1, 10)	(0.03, 45)	(0.026, 10)	(0.1, 10)	(0.03, 45)
TP % (sensitivity)	75 (36/48)	83.3 (40/48)	75 (36/48)	75 (36/48)	85.4 (41/48)	70.8 (34/48)	75 (36/48)	75 (36/48)	54.2 (26/48)
TN % (specificity)	63.6 (14/22)	45.5 (10/22)	63.6 (14/22)	50 (11/22)	45.5 (10/22)	59.1 (13/22)	63.6 (14/22)	63.6 (14/22)	68.2 (15/22)
FP %	36.4 (8/22)	54.5 (12/22)	36.4 (8/22)	50 (11/22)	54.5 (12/22)	40.9 (9/22)	36.4 (8/22)	36.4 (8/22)	31.8 (7/22)
FN %	25 (12/48)	16.7 (8/48)	25 (12/48)	25 (12/48)	14.6 (7/48)	29.2 (14/48)	25 (12/48)	25 (12/48)	45.8 (22/48)
Average Accuracy	0.693	0.644	0.693	0.625	0.654	0.65	0.693	0.693	0.612
Overall Accuracy	0.714	0.714	0.714	0.671	0.729	0.671	0.714	0.714	0.586

Table 4-3: Model 3 (I_{1u}/IC_{50} or I_2/IC_{50})

Model 3 (I_{1u}/IC_{50} or I_2/IC_{50})	IC_{50}(V)[a]		IC_{50}(ER)[a]		IC_{50}(NSF)[a]	
	ROC Analysis	EMA DDI Guidance[e]	ROC Analysis	EMA DDI Guidance[e]	ROC Analysis	EMA DDI Guidance[e]
Cut-off values	(0.00141, 3334)	(0.02,10)	(0.00177,62)	(0.02,10)	(0.00052,94)	(0.02,10)
TP % (sensitivity)	75 (36/48)	83.3 (40/48)	75 (36/48)	85.4 (41/48)	75 (36/48)	75 (36/48)
TN % (specificity)	81.8 (18/22)	45.5 (10/22)	54.5 (12/22)	45.5 (10/22)	68.2 (15/22)	63.6 (14/22)
FP %	18.2 (4/22)	54.5 (12/22)	45.5 (10/22)	54.5 (12/22)	31.8 (7/2 + 2)	36.4 (8/22)
FN %	25 (12/48)	16.7 (8/48)	25 (12/48)	14.6 (7/48)	25 (12/48)	25 (12/48)
Average Accuracy	0.784	0.644	0.648	0.654	0.716	0.693
Overall Accuracy	0.771	0.714	0.686	0.729	0.729	0.714

[a]IC_{50}(V), IC_{50}(ER), and IC_{50}(NSF) were used in the static models for DDI prediction, respectively. IC_{50}(V), IC_{50}(ER), and IC_{50}(NSF) were determined as described in the *Materials and Methods* section and shown in Table 3.
[b]The cut-off value was obtained from FDA final DDI guidance (FDA, 2020).
[c]The cut-off value was obtained from Ellens et al. (2013) based on ROC analysis of P-gp IC_{50} data for 15 compounds generated by 23 laboratories using four in vitro systems: Caco-2 cells, LLCPK1-MDR1, MDCKII-MDR1, and MDR1 vesicles. P-gp probe substrates were digoxin for polarized cell-lines and NMQ or vinblastine for MDR1 vesicles.
[d]The cut-off value was obtained from FDA DDI draft guidance (Agarwal et al., 2013).
[e]The cut-off value was obtained from EMA DDI guidance EMA (2012).

et al. (2013)]. In addition, the cut-off and predictive performance using vesicular IC_{50} data generated in a single laboratory [laboratory 20, Supplemental Table 2 of Ellens et al. (2013)] was compared. In brief, the cut-off value ($I_1/IC_{50} \geq 0.08$, $I_2/IC_{50} \geq 501$) derived from all MDR1 vesicular IC_{50} data by Ellens et al. (2013) resulted in a FN% and FP% of 33% and 8%, respectively (FN% and FP % of 36% and 33%, respectively in a test set). The cut-off value ($I_1/IC_{50} \geq 0.01$, $I_2/IC_{50} \geq 89$) derived from vesicular IC_{50} data in a single laboratory (laboratory 20) showed a FN % and FP % of 36% and 18%, respectively. In contrast, the cut-off value derived from our vesicular data ($I_1/IC_{50} \geq 0.032$, $I_2/IC_{50} \geq 40$; Table 4-2) showed lower FN% (25%), but higher FP% (36.4%). The difference in the cut-off values derived from different laboratories may be attributed to interlaboratory variability of IC_{50} data measured with MDR1 vesicles and a different set of training compounds used in the calibration. This highlights the need to calibrate MDR1 vesicular assays for P-gp DDI prediction. The difference in probe substrates (NMQ versus vinblastine), the source of membrane vesicles (P-gp expression levels, the ratio of inside-out to right-side-out vesicles), and assay conditions may contribute to such interlaboratory variability.

In model 3 (I_{1u}/IC_{50}, I_2/IC_{50}), the cut-off values derived based on IC_{50}(V), IC_{50}(ER), and IC_{50}(NSF) was (0.00141, 3334), (0.00177, 62), and (0.00052, 94), respectively, which was approximately 12–38 lower in I_{1u}/IC_{50}, but 6–333-fold higher in I_2/IC_{50} than the threshold recommended by EMA ($I_{1u}/IC_{50} \geq 0.02$ or $I_2/IC_{50} \geq 10$)

(EMA, 2012). Under the predefined sensitivity of 0.75 (FN% 25%), ROC refined cut-off value with IC_{50}(V) resulted in the lowest FP% (18.2%) compared to that of IC_{50}(ER) (45.5%) and IC_{50}(NSF) (31.8%). EMA cut-off value generally resulted in lower FN% (14.6%–25%), but higher FP% (36.4%–54.5%). The list of FN and FP compounds in this training set under different models and cut-off values was summarized in Supplemental Table 5.

A total of 12 compounds were selected as an independent test set to evaluate the performance of ROC refined cut-offs using IC_{50}(V), IC_{50}(ER), and IC_{50}(NSF) (Supplemental Table 6) and clinical digoxin (oral dose only) DDI data (Supplemental Table 2; 8 positive, 4 negative DDIs). The predictive performance of the test set was summarized in Supplemental Table 7. Within this limited data set, for I_2/IC_{50} model, ROC derived cut-off values for both IC_{50}(V), and IC_{50}(NSF) resulted in no FN, whereas IC_{50}(ER) showed 12.5% FN. The cut-off values by FDA and Ellens et al. (2013) provided reasonably good performance but higher FN% for the cut-off by Ellens et al. (2013). Similar results were shown for model 2 (I_1/IC_{50} or I_2/IC_{50}): IC_{50}(V) and IC_{50}(NSF) did not show FN, whereas IC_{50}(ER) had FN% of 12.5. For model 3 (I_{1u}/IC_{50}, I_2/IC_{50}), a higher FN% for ROC derived cut-offs than the one recommended by EMA was observed for IC_{50}(V), IC_{50}(ER), and IC_{50}(NSF). The list of FN and FP prediction of the compounds in the test set was summarized in Supplemental Table 8.

Prediction of P-gp-Mediated DDIs for Selected Compounds Using DE and Fexofenadine as Clinical Probe Drugs

To understand whether in vitro P-gp IC_{50} values measured using digoxin and NMQ as probes and the cut-off values derived from digoxin clinical DDI data can be used to predict DDIs for other P-gp probe substrates, we extended our analysis by evaluating predictive performance of 23 P-gp related clinical DDI data using DE (17) and fexofenadine (6) as probes (Table 2, 19 positive 4 negative DDIs, respectively) for 15 inhibitor drugs with IC_{50} values measured in our studies (Table 3). Only the performance of model 1 (I_2/IC_{50}) was evaluated, as gut P-gp is the major site for P-gp related DDIs with DE and fexofenadine (Lappin et al., 2010; Li et al., 2017; Chu et al., 2018a, 2018b). In this limited dataset, applying the cut-off values derived from ROC analysis based on clinical digoxin DDI data, the FN% of DDI prediction for DE and fexofenadine using IC_{50}(V), IC_{50}(ER), and IC_{50}(NSF) was 15.8%, 10.5%, and 0%, respectively, whereas the FP% was 50%, 50%, and 75%, respectively (Table 5). Likewise, FDA cut-off value resulted in low FN%, but high FP%. The cut-off value by Ellens et al. (2013) using IC_{50}(V) and IC_{50}(ER) showed FN% of 21.1% and 26.3%, and FP% of 25% and 25%, respectively, whereas IC_{50}(NSF) data resulted in high FN% (52.6%). The list of compounds with FN and FP prediction in this dataset was summarized in Supplemental Table 9. In model 1, three FN predictions were observed using IC_{50}(V). However, the magnitude of clinical DDIs ranged only 1.26- to 1.68-fold, indicating weak to moderate DDIs. Two FP predictions using IC_{50}(V) were atorvastatin and ritonavir, BCS Class II and IV compound, respectively, which is likely caused by lower inhibitor concentration in the gut due to low solubility.

Discussion

To improve the prediction of P-gp related DDIs, we have systemically calibrated our P-gp inhibition assays (LLC-MDR1 cells and MDR1 vesicles). This is for the first time a side-by-side comparison of two P-gp inhibition assays conducted in the same laboratory. To our knowledge, the numbers of compounds calibrated in MDR1 vesicles in this study are larger than those reported in the literature to date (Ellens et al., 2013; Heredi-Szabo et al., 2013; Fekete et al., 2015).

There is a system-dependent difference of IC_{50} values between BDT and vesicular inhibition assays. Despite a limited dataset, it appears that more difference in IC_{50}s was observed for BCS class II and IV compounds. For instance, IC_{50}(NSF), IC_{50}(ER), and IC_{50}(UDF) of velpatasvir and daclatasvir were more than 10- and 20-fold higher than IC_{50}(V) and resulted in false negative DDI predictions. IC_{50}(V) of fidaxomicin was at least 224-fold lower than IC_{50}(NSF), IC_{50}(ER), and IC_{50}(UDF). The correction for nonspecific binding in BDT assay did not reduce such difference. In contrast, a FP prediction was observed for fidaxomicin using IC_{50}(V). The mechanisms for such difference are not known, likely caused by substrate-dependent inhibition (digoxin versus NMQ). Furthermore, in inside-out membrane vesicles, inhibitor drugs have direct access to P-gp binding sites, whereas in BDT assay, inhibitors need to permeate across the lipid bilayers to access P-gp binding sites located in the cytosolic leaflet of the plasma membrane. This could yield a difference of apparent IC_{50} values for a poorly permeable inhibitor drug, like fidaxomicin. For BDT assay, there were also some differences in IC_{50} values generated by different calculation methods (ER, NSF, and UDF) even if the same dataset was used. However, such difference was generally less profound than those between BDT and vesicular inhibition assay. Currently, there is no consensus on the optimal method to calculate P-gp IC_{50} in BDT assay. The model-based approach may provide a more mechanistic and accurate estimation of IC_{50} and K_i values (Kishimoto et al., 2016).

The predictive performances of various static models, IC_{50} measurement, and calculation methods were compared. Among 3 static models

TABLE 5

Summary of predictive performance of selected P-gp clinical inhibition studies using DE and fexofenadine as in vivo probes and the cut-off values derived from ROC analysis based on digoxin clinical DDI data and the comparison with other cut-off criteria

	IC_{50}(V)[a]			IC_{50}(ER)[a]			IC_{50}(NSF)[a]		
Model 1 (I_2/IC_{50})	ROC Analysis[b]	FDA , 2020[c]	Ellens et al., 2013[d]	ROC Analysis[b]	FDA , 2020[c]	Ellens et al., 2013[d]	ROC Analysis[b]	FDA , 2020[c]	Ellens et al., 2013[d]
Cut-off values	25.9	10.0	45.0	13.7	10.0	45.0	9.3	10.0	45.0
TP % (sensitivity)	84.2 (16/19)	94.7 (18/19)	78.9 (15/19)	89.5 (17/19)	94.7 (18/19)	73.7 (14/19)	100 (19/19)	100 (19/19)	47.4 (9/19)
TN % (specificity)	50 (2/4)	25 (1/4)	75 (3/4)	50 (2/4)	50 (2/4)	75 (3/4)	25 (1/4)	50 (2/4)	100 (4/4)
FP %	50 (2/4)	75 (3/4)	25 (1/4)	50 (2/4)	50 (2/4)	25 (1/4)	75 (3/4)	50 (2/4)	0 (0/4)
FN %	15.8 (3/19)	5.3 (1/19)	21.1 (4/19)	10.5 (2/19)	5.3 (1/19)	26.3 (5/19)	0 (0/19)	0 (0/19)	52.6 (10/19)
Average Accuracy	0.783	0.826	0.783	0.826	0.870	0.739	0.870	0.913	0.565
Overall Accuracy	0.671	0.599	0.770	0.697	0.724	0.743	0.625	0.750	0.737

[a]IC_{50}(NSF), IC_{50}(ER), and IC_{50}(V) were used in the static models for DDI prediction, respectively. IC_{50}(NSF), IC_{50}(ER), and IC_{50}(V) of respective inhibitor drugs were shown in Table 3.
[b]Cut-off values derived from ROC analysis based on digoxin clinical DDI data in Table 1 and IC_{50} data in Table 3.
[c]The cut-off value was obtained from FDA final DDI guidance (FDA, 2020).
[d]The cut-off value was obtained from Ellens et al. (2013) based on ROC analysis of P-gp IC_{50} data for 15 compounds generated by 23 laboratories using four in vitro systems: Caco-2 cells, LLCPK1-MDR1, MDCKII-MDR1, and MDR1 vesicles. P-gp probe substrates were digoxin for polarized cell-lines and NMQ or vinblastine for MDR1 vesicles.

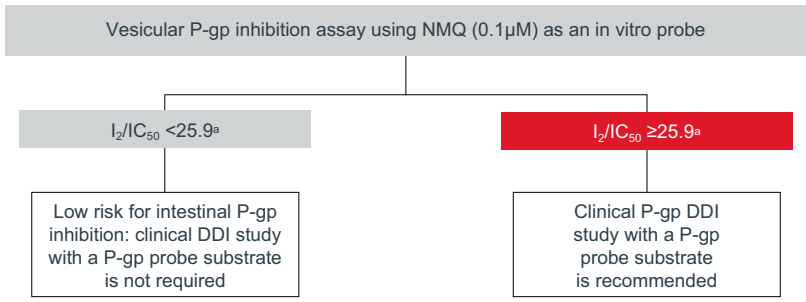

Fig. 5. Recommended workflow to evaluate intestinal P-gp inhibition. This workflow is only suitable when orally administered digoxin is used as a clinical P-gp probe. Additional calibration will be needed for dabigatran etexilate (DE) and other orally administered P-gp probe substrates. If digoxin is a comedication, IC_{50}(NSF) or IC_{50}(ER) measured by BDT assay using digoxin as a probe substrate may be helpful in clinical study design and data interpretation. [a]:The threshold of 25.9 is obtained based on calibration of vesicular P-gp inhibition assay using ROC analysis shown in this paper. The calibration of your own assay using similar approach is recommended.

evaluated, inclusion of I_1/IC_{50}(model 2) or I_{1u}/IC_{50}(model 3) did not provide superior predictive performance over model 1(I_2/IC_{50} only) based on ROC_{AUC}, FN%, FP%, and the accuracy of DDI prediction. Furthermore, different from model 1, 2D ROC curves in models 2 and 3 were in zigzig shape, suggesting that the standard ROC analysis for 2D classifiers was not well defined and might need further statistical modeling of the 2D predictors to follow suit. This may also be attributed to the fact that only orally administered digoxin clinical DDI data were collected in this analysis. This observation, similar to others (Poirier et al., 2014; Zhou et al., 2019), suggested that inhibition of intestinal, but not renal P-gp is the major DDI mechanism for orally administered digoxin. Therefore, model 1 (I_2/IC_{50}) is sufficient to predict DDIs for orally administered P-gp substrates for simplistic and practical considerations, whereas models 2 and 3 did not improve prediction accuracy, but rather add uncertainty for highly bound inhibitor drugs when f_u cannot be accurately measured.

Based on $AUC_{ROC,}$ IC_{50}(V) appeared to show the trend of better predictive performance, followed by IC_{50}(ER), IC_{50}(NSF), and IC_{50}(UDF) across all 3 static models evaluated, despite the lack of statistical significance, due to the limited sample sizes. Overall, the optimal discrimination thresholds derived from this training set for respective in vitro IC_{50} data had a minimal sensitivity of 0.75, and highest specificity. For example, in model 1 (I_2/IC_{50}), IC_{50}(V), IC_{50}(ER), and IC_{50}(NSF) had similar accuracy values. IC_{50}(ER) demonstrated lowest FN%, but highest FP%, whereas IC_{50}(V) and IC_{50}(NSF) showed comparable FN% and FP%. Compared with the cut-off values recommended by FDA and Ellens et al. (2013), the cut-offs derived from our ROC analysis exhibited better or comparable predictive performance, which highlighted the need to calibrate in vitro systems to provide more accurate DDI prediction. The static models and ROC derived cut-off values were further validated with an independent test set of 12 compounds, confirming a good predictive performance.

In training set, there were a total of 12 FN prediction for IC_{50}(V) using optimized I_2/IC_{50} cut off value (25.9). For these FN predictions, 5 out of 12 had AUCR or $C_{max}R < 1.5$, suggesting weak DDIs, whereas 7 out of 12 had AUCR or $C_{max}R$ ranged 1.5–2. For these 7 studies, the mechanisms of underprediction were not well understood. For captopril, FN prediction was observed in all models and assays using either optimized cut-offs or values recommended by FDA, EMA, and Ellens et al. (2013). Carvedilol underprediction was only observed in DDI studies with female subjects. Underprediction of flibanserin DDIs could be attributed to substrate- and/or system- dependent difference of IC_{50} measurement, as FN prediction was not shown using IC_{50}(NSF) and IC_{50}(ER).

As digoxin is neither a specific nor sensitive P-gp probe (Taub et al., 2011; Lee et al., 2014), DE has been recommended by regulatory agencies as an alternative clinical probe for gut P-gp inhibition (FDA, 2020; EMA, 2012). However, DE was neither stable in cell-based assays (Chu et al., 2018a), nor showing robust transport in MDR1 vesicles (unpublished observations) likely due to higher P_{app} and nonspecific binding. This has precluded the use of DE as an in vitro probe for P-gp inhibition. In this study, we explored the feasibility of using NMQ and digoxin and the cut-offs derived from digoxin clinical DDI studies (model 1) to predict P-gp related DDIs for DE and fexofenadine, another P-gp substrate (Chu et al., 2018b). Due to limited in vitro and clinical data, only 23 clinical studies were selected in this analysis. Overall, a lower FN% was observed for IC_{50}(V) and IC_{50}(ER) with higher FP%. As only a total of 4 negative DDI data were available in this dataset, it may not be feasible to accurately assess the predictive performance, especially for specificity. More compounds should be included in future studies.

Beyond high IC_{50} variability, there are several issues and knowledge gaps for P-gp DDI prediction. 1) Prediction of DDIs for perpetrators with low solubility. For class II/IV compounds, the solubility of inhibitor drugs can be much lower than I_2. I_2/IC_{50} model may not be feasible to derisk DDIs due to the inability to test the inhibitor concentration beyond the solubility limit. In this case, we assumed the highest concentrations tested as surrogate IC_{50}. Such estimation may introduce additional variability, if different highest inhibitor concentrations are tested. It also cannot differentiate non versus weak inhibition and makes the prediction not definitive. Furthermore, DDIs may be overestimated if the inhibitor concentration in gut is lower than I_2. This is consistent with our observations that 11 out of 15 FP prediction (model 1) are for class II or IV compounds (Supplemental Table 5). In the future, prediction of inhibitor concentration in gut using mechanistic modeling may improve the accuracy of DDI prediction. 2) Relevant inhibitor concentrations for IC_{50} measurement may be different from nominal concentrations. However, correction of nonspecific binding of inhibitor drugs in BDT assay did not improve the predictive performance. It is possible that unbound inhibitor concentrations measured in the incubation medium are not relevant concentrations for P-gp inhibition, as substrate binding sites of P-gp are localized intracellularly. Therefore, measuring intracellular unbound inhibitor concentrations and developing in vitro mechanistic modeling to determine true K_i values may improve the prediction of DDIs. 3) Mechanisms for P-gp inhibition have not been well characterized. It is not clear whether it is driven by cis- and/or transinhibition. Understanding such mechanisms will help to improve in vitro to in vivo extrapolation. 4) It is worth noting that possible involvement of other mechanisms for the disposition of P-gp probe drugs, e.g., digoxin, DE, and fexofenadine (Shimizu et al., 2005; Taub et al., 2011; Chu et al., 2018a;

Medwid et al., 2019; Shen et al., 2019), may confound the DDI prediction, as they are not captured in such simplified P-gp inhibition models.

In conclusion, our studies confirmed that model 1 (I_2/IC_{50}) is sufficient to predict P-gp mediated intestinal DDIs. In evaluation of 7 P-gp IC_{50} measurement/calculation methods, IC_{50}(V), IC_{50}(NSF), and IC_{50}(ER) provided better predictive performance than others. Further analysis indicated that IC_{50}(V) obtained from MDR1 vesicles with refined threshold of $I_2/IC_{50} \geq 25.9$ provided comparable predictive power over BDT assays in LLC-MDR1 cells with IC_{50}(NSF) and IC_{50}(ER) at the threshold of $I_2/IC_{50} \geq 9.3$ and 13.7, respectively. These IC_{50} assays and cut-off values could also be used to predict P-gp mediated intestinal DDIs for DE and fexofenadine, despite that additional data are needed for further validation. Through comprehensive calibration, our studies demonstrate that IC_{50} data generated from vesicular inhibition assay using NMQ as probe substrate are predictive for P-gp related intestinal DDIs with digoxin. We therefore recommend vesicular inhibition assay as our preferred method to study P-gp-mediated intestinal DDIs for digoxin (Fig. 5), due to its simplicity, lower variability, higher assay throughput, and more direct estimation of kinetic parameters over BDT assays. This assay also offers the potential to predict DDIs for other P-gp probe substrates provided confirmatory validation is conducted.

Acknowledgments

The authors are thankful to Lisa G. Fox, Elke Perloff, and J. George Zhang from Corning Life Sciences for their expert assistance with experiments.

Authorship Contributions

Participated in research design: Yabut, Houle, Wang, Liaw, Chu.

Conducted experiments: Yabut, Houle, Katwaru.

Performed data analysis: Yabut, Houle, Wang, Collier, Chu.

Wrote or contributed to the writing of the manuscript: Yabut, Houle, Wang, Liaw, Katwaru, Hittle, Chu.

References

Agarwal S, Arya V, and Zhang L (2013) Review of P-gp inhibition data in recently approved new drug applications: utility of the proposed [I(1)]/IC(50) and [I(2)]/IC(50) criteria in the P-gp decision tree. *J Clin Pharmacol* **53**:228–233.

Balimane PV, Marino A, and Chong S (2008) P-gp inhibition potential in cell-based models: which "calculation" method is the most accurate? *AAPS J* **10**:577–586.

Benet LZ (2013) The role of BCS (biopharmaceutics classification system) and BDDCS (biopharmaceutics drug disposition classification system) in drug development. *J Pharm Sci* **102**:34–42.

Bentz J, O'Connor MP, Bednarczyk D, Coleman J, Lee C, Palm J, Pak YA, Perloff ES, Reyner E, Balimane P, et al. (2013) Variability in P-glycoprotein inhibitory potency (IC_{50}) using various in vitro experimental systems: implications for universal digoxin drug-drug interaction risk assessment decision criteria. *Drug Metab Dispos* **41**:1347–1366.

Brouwer KL, Keppler D, Hoffmaster KA, Bow DA, Cheng Y, Lai Y, Palm JE, Stieger B, and Evers R; International Transporter Consortium (2013) In vitro methods to support transporter evaluation in drug discovery and development. *Clin Pharmacol Ther* **94**:95–112.

Chan G, Houle R, Lin M, Yabut J, Cox K, Wu J, and Chu X (2019) Role of transporters in the disposition of a novel β-lactamase inhibitor: relebactam (MK-7655). *J Antimicrob Chemother* **74**:1894–1903.

Chu X, Galetin A, Zamek-Gliszczynski MJ, Zhang L, and Tweedie DJ; International Transporter Consortium (2018a) Dabigatran etexilate and digoxin: comparison as clinical probe substrates for evaluation of P-gp inhibition. *Clin Pharmacol Ther* **104**:788–792.

Chu X, Liao M, Shen H, Yoshida K, Zur AA, Arya V, Galetin A, Giacomini KM, Hanna I, Kusuhara H, et al.; International Transporter Consortium (2018b) Clinical probes and endogenous biomarkers as substrates for transporter drug-drug interaction evaluation: perspectives from the international transporter consortium. *Clin Pharmacol Ther* **104**:836–864.

Cook JA, Feng B, Fenner KS, Kempshall S, Liu R, Rotter C, Smith DA, Troutman MD, Ullah M, and Lee CA (2010) Refining the in vitro and in vivo critical parameters for P-glycoprotein, [I]/IC50 and [I2]/IC50, that allow for the exclusion of drug candidates from clinical digoxin interaction studies. *Mol Pharm* **7**:398–411.

Ellens H, Deng S, Coleman J, Bentz J, Taub ME, Ragueneau-Majlessi I, Chung SP, Herédi-Szabó K, Neuhoff S, Palm J, et al. (2013) Application of receiver operating characteristic analysis to refine the prediction of potential digoxin drug interactions. *Drug Metab Dispos* **41**:1367–1374.

FDA (2020) In Vitro Drug Interaction Studies — Cytochrome P450 Enzyme- and Transporter-Mediated Drug Interactions Guidance for Industry, in *Center for Drug Evaluation and Research (CDER)*.

Fekete Z, Rajnai Z, Nagy T, Jakab KT, Kurunczi A, Gémes K, Herédi-Szabó K, Fülöp F, Tóth GK, Czerwinski M, et al. (2015) Membrane assays to characterize interaction of drugs with ABCB1. *J Membr Biol* **248**:967–977.

Fenner KS, Troutman MD, Kempshall S, Cook JA, Ware JA, Smith DA, and Lee CA (2009) Drug-drug interactions mediated through P-glycoprotein: clinical relevance and in vitro-in vivo correlation using digoxin as a probe drug. *Clin Pharmacol Ther* **85**:173–181.

Giacomini KM, Huang SM, Tweedie DJ, Benet LZ, Brouwer KL, Chu X, Dahlin A, Evers R, Fischer V, Hillgren KM, et al.; International Transporter Consortium (2010) Membrane transporters in drug development. *Nat Rev Drug Discov* **9**:215–236.

Hanley JA and McNeil BJ (1982) The meaning and use of the area under a receiver operating characteristic (ROC) curve. *Radiology* **143**:29–36.

Herédi-Szabó K, Palm JE, Andersson TB, Pál Á, Méhn D, Fekete Z, Beéry E, Jakab KT, Jani M, and Krajcsi P (2013) A P-gp vesicular transport inhibition assay - optimization and validation for drug-drug interaction testing. *Eur J Pharm Sci* **49**:773–781.

Jani M and Krajcsi P (2014) In vitro methods in drug transporter interaction assessment. *Drug Discov Today Technol* **12**:e105–e112.

Kishimoto W, Ishiguro N, Ludwig-Schwellinger E, Ebner T, Maeda K, and Sugiyama Y (2016) Usefulness of a model-based approach for estimating in vitro P-glycoprotein inhibition potency in a transcellular transport assay. *J Pharm Sci* **105**:891–896.

Lappin G, Shishikura Y, Jochemsen R, Weaver RJ, Gesson C, Houston B, Oosterhuis B, Bjerrum OJ, Rowland M, and Garner C (2010) Pharmacokinetics of fexofenadine: evaluation of a microdose and assessment of absolute oral bioavailability. *Eur J Pharm Sci* **40**:125–131.

Lee CA, Cook JA, Reyner EL, and Smith DA (2010) P-glycoprotein related drug interactions: clinical importance and a consideration of disease states. *Expert Opin Drug Metab Toxicol* **6**:603–619.

Lee CA, Kalvass JC, Galetin A, and Zamek-Gliszczynski MJ (2014) ITC commentary on the prediction of digoxin clinical drug-drug interactions from in vitro transporter assays. *Clin Pharmacol Ther* **96**:298–301.

Li F, Howard KD, and Myers MJ (2017) Influence of P-glycoprotein on the disposition of fexofenadine and its enantiomers. *J Pharm Pharmacol* **69**:274–284.

Medwid S, Li MMJ, Knauer MJ, Lin K, Mansell SE, Schmerk CL, Zhu C, Griffin KE, Yousif MD, Dresser GK, et al. (2019) Fexofenadine and rosuvastatin pharmacokinetics in mice with targeted disruption of organic anion transporting polypeptide 2B1. *Drug Metab Dispos* **47**:832–842.

Nervi P, Li-Blatter X, Aänismaa P, and Seelig A (2010) P-glycoprotein substrate transport assessed by comparing cellular and vesicular ATPase activity. *Biochim Biophys Acta* **1798**:515–525.

O'Connor M, Lee C, Ellens H, and Bentz J (2015) A novel application of t-statistics to objectively assess the quality of IC50 fits for P-glycoprotein and other transporters. *Pharmacol Res Perspect* **3**:e00078.

Papich MG and Martinez MN (2015) Applying biopharmaceutical classification system (BCS) criteria to predict oral absorption of drugs in dogs: challenges and pitfalls. *AAPS J* **17**:948–964.

Poirier A, Cascais AC, Bader U, Portmann R, Brun ME, Walter I, Hillebrecht A, Ullah M, and Funk C (2014) Calibration of in vitro multidrug resistance protein 1 substrate and inhibition assays as a basis to support the prediction of clinically relevant interactions in vivo. *Drug Metab Dispos* **42**:1411–1422.

Robin X, Turck N, Hainard A, Tiberti N, Lisacek F, Sanchez JC, and Müller M (2011) pROC: an open-source package for R and S+ to analyze and compare ROC curves. *BMC Bioinformatics* **12**:77.

Saito H, Osumi M, Hirano H, Shin W, Nakamura R, and Ishikawa T (2009) Technical pitfalls and improvements for high-speed screening and QSAR analysis to predict inhibitors of the human bile salt export pump (ABCB11/BSEP). *AAPS J* **11**:581–589.

Shen H, Yao M, Sinz M, Marathe P, Rodrigues AD, and Zhu M (2019) Renal excretion of dabigatran: the potential role of multidrug and toxin extrusion (MATE) proteins. *Mol Pharm* **16**:4065–4076.

Shimizu M, Fuse K, Okudaira K, Nishigaki R, Maeda K, Kusuhara H, and Sugiyama Y (2005) Contribution of OATP (organic anion-transporting polypeptide) family transporters to the hepatic uptake of fexofenadine in humans. *Drug Metab Dispos* **33**:1477–1481.

Sugimoto H, Matsumoto S, Tachibana M, Niwa S, Hirabayashi H, Amano N, and Moriwaki T (2011) Establishment of in vitro P-glycoprotein inhibition assay and its exclusion criteria to assess the risk of drug-drug interaction at the drug discovery stage. *J Pharm Sci* **100**:4013–4023.

Sziráki I, Erdo F, Beéry E, Molnár PM, Fazakas C, Wilhelm I, Makai I, Kis E, Herédi-Szabó K, Abonyi T, et al. (2011) Quinidine as an ABCB1 probe for testing drug interactions at the blood-brain barrier: an in vitro in vivo correlation study. *J Biomol Screen* **16**:886–894.

Taub ME, Mease K, Sane RS, Watson CA, Chen L, Ellens H, Hirakawa B, Reyner EL, Jani M, and Lee CA (2011) Digoxin is not a substrate for organic anion-transporting polypeptide transporters OATP1A2, OATP1B1, OATP1B3, and OATP2B1 but is a substrate for a sodium-dependent transporter expressed in HEK293 cells. *Drug Metab Dispos* **39**:2093–2102.

Volpe DA, Hamed SS, and Zhang LK (2014) Use of different parameters and equations for calculation of IC_{50} values in efflux assays: potential sources of variability in IC_{50} determination. *AAPS J* **16**:172–180.

Wu CY and Benet LZ (2005) Predicting drug disposition via application of BCS: transport/absorption/ elimination interplay and development of a biopharmaceutics drug disposition classification system. *Pharm Res* **22**:11–23.

Zamek-Gliszczynski MJ, Lee CA, Poirier A, Bentz J, Chu X, Ellens H, Ishikawa T, Jamei M, Kalvass JC, Nagar S, et al.; International Transporter Consortium (2013) ITC recommendations for transporter kinetic parameter estimation and translational modeling of transport-mediated PK and DDIs in humans. *Clin Pharmacol Ther* **94**:64–79.

Zhang L, Huang SM, Reynolds K, Madabushi R, and Zineh I (2018) Transporters in drug development: scientific and regulatory considerations. *Clin Pharmacol Ther* **104**:793–796.

Zhou T, Arya V, and Zhang L (2019) Comparing various in vitro prediction methods to assess the potential of a drug to inhibit P-glycoprotein (P-gp) transporter in vivo. *J Clin Pharmacol* **59**:1049–1060.

Address correspondence to: Dr. Xiaoyan Chu, Merck & Co. Inc., RY800-D211, 126 East Lincoln Avenue, Rahway, NJ 07065. E-mail: xiaoyan_chu@merck.com

Selection of an optimal in vitro model to assess P-gp inhibition: comparison of vesicular and bi-directional transcellular transport inhibition assays

Jocelyn Yabut*, Robert Houle*, Shubing Wang, Andy Liaw, Ravi Katwaru, Hannah Collier, Lucinda Hittle, and Xiaoyan Chu

Department of Pharmacokinetics, Pharmacodynamics and Drug Metabolism (JY, RH, RK, HC, LH, and XC), and Department of Biometrics Research (SW and AL), Merck & Co., Inc., Kenilworth, New Jersey, USA

*: JY and RH contributed equally to this work

Running title: In vitro models to study P-gp-mediated intestinal DDIs

Address correspondence to:

Xiaoyan Chu, Ph.D.

Merck & Co. Inc., RY800-D211

126 East Lincoln Avenue

Rahway, NJ 07065, USA

Tel 732-594-0977; Fax 732-594-2382; e-mail xiaoyan_chu@merck.com

Number of Text Pages: 36

Number of Tables: 5

Number of Figures: 5

Number of references: 41

Number of words in Abstract: 248

Number of words in Introduction: 750

Number of words in Discussion: 1499

Nonstandard abbreviations used:

AUC, area under the curve; AUC_{ROC}, area under the ROC curve; C_{max}, maximum concentration; DDI, drug-drug interaction; DE, dabigatran etexilate; EMA, European Medical Agency; ER, efflux ratio; FDA, Food and Drug Administration; FN, false negatives; FP, false positives; FPR, false positive rate; $[I_1]$, maximum concentration of inhibitor at steady state; $[I_{1u}]$, unbound $[I_1]$; $[I_2]$, concentration of inhibitor in the gastrointestinal tract based on highest approved dose dissolved by 250 ml; IC_{50}, half-maximal inhibitory concentration; LLC-PK1-MDR1, Lilly Laboratories Cell Porcine Kidney 1 cells overexpressing MDR1; MDR1, multidrug resistance protein 1; NME, new molecular entity; P_{app}, apparent permeability; P-gp, P-glycoprotein; PK, pharmacokinetic; ROC, receiver operating characteristic; TN, true negatives; TP, true positives; TPR, true positive rate.

Abstract

The multidrug resistance protein 1 (MDR1) P-glycoprotein (P-gp) is a clinically important transporter. In vitro P-gp inhibition assays have been routinely conducted to predict the potential for clinical drug-drug interactions (DDIs) mediated by P-gp. However, high inter- laboratory and inter-system variability of P-gp IC_{50} data limits accurate prediction of DDIs using static models and decision criteria recommended by regulatory agencies. In this study, we calibrated two in vitro P-gp inhibition models: vesicular uptake of N-methyl-quinidine (NMQ) in MDR1 vesicles and bidirectional transport (BDT) of digoxin in Lilly Laboratories Cell Porcine Kidney 1 cells overexpressing MDR1 (LLC-MDR1) using a total of 48 P-gp inhibitor and non-inhibitor drugs, and digoxin DDI data from 70 clinical studies. Refined thresholds were derived using receiver operating characteristic (ROC) analysis and their predictive performance was compared with the decision frameworks proposed by regulatory agencies and selected reference. Furthermore, the impact of various IC_{50} calculation methods and non-specific binding of drugs on DDI prediction was evaluated. Our studies suggest that the concentration of inhibitor based on highest approved dose dissolved in 250 ml divided by IC_{50}(I_2/IC_{50}) is sufficient to predict P-gp related intestinal DDIs. IC_{50} obtained from vesicular inhibition assay with a refined threshold of $I_2/IC_{50} \geq 25.9$ provides comparable predictive power than those measured by net secretory flux and efflux ratio in LLC-MDR1 cells. We therefore recommend vesicular P-gp inhibition as our preferred method given its simplicity, lower variability, higher assay throughput, and more direct estimation of in vitro kinetic parameters than BDT assay.

Significance Statement

We have conducted comprehensive calibration of two in vitro P-gp inhibition models: uptake in MDR1 vesicles and bidirectional transport in LLC-MDR1 cell monolayers to predict DDIs. Our studies suggest that IC_{50}s obtained from vesicular inhibition with a refined threshold of $I_2/IC_{50} \geq 25.9$ provide comparable predictive power than those in LLC-MDR1 cells. We therefore recommend vesicular P-gp inhibition as preferred method given its simplicity, lower variability, higher assay throughput, and more direct estimation of in vitro kinetic parameters.

Introduction

Multidrug resistance protein 1 (MDR1) P-glycoprotein (P-gp) is a clinically important transporter (International Transporter et al., 2010; Lee et al., 2010). Inhibition of P-gp can cause drug–drug interactions (DDIs), in which, inhibition of intestinal P-gp appears to have the most significant impact (Fenner et al., 2009; International Transporter et al., 2010; Zhang et al., 2018). Thus, the US Food and Drug Administration (FDA) and the European Medicines Agency (EMA) have requested evaluating the potential of a new molecular entity (NME) to inhibit P-gp in vitro and recommended a decision framework to determine the need for conducting clinical DDI studies with P-gp probe substrates, such as digoxin and dabigatran etexilate (DE) (EMA, 2012; FDA, 2020).

In vitro P-gp inhibition studies are routinely conducted in the pharmaceutical industry to evaluate the potential of NMEs as in vivo inhibitors of P-gp based on the recommendations from regulatory agencies. To assess their predictability for digoxin DDIs, a P-gp IC_{50} working group measured in vitro P-gp IC_{50} values for 15 compounds in 23 laboratories using their own assays and protocols. Substantial inter-laboratory variability for IC_{50} values was reported (Bentz et al., 2013). A receiver operating characteristic (ROC) analysis was conducted by this group and refined cut-off values were proposed, which accounted for inter-laboratory variability for IC_{50} values (Ellens et al., 2013). Currently, P-gp IC_{50} variability is still a major concern which precludes accurate DDI prediction. As such, EMA has recommended to assess P-gp inhibition using two separate in vitro systems. A systemic calibration of in vitro assays in each individual laboratory may help to address this issue before the standardized model and assay protocol are

established and employed. For instance, Cook et al (Cook et al., 2010), Sugimoto et al (Sugimoto et al., 2011), and Poirier et al (Poirier et al., 2014) calibrated their P-gp inhibition assays in Caco-2 cells (human colon adenocarcinoma cells expressing endogenous P-gp) and LLC-MDR1 cells (Lilly Laboratories Cell Porcine Kidney 1 cells overexpressing MDR1) using 26 to 68 clinical digoxin DDI data and defined their cut-off criteria accordingly.

In vitro P-gp inhibition is most frequently evaluated in P-gp transfected cells and Caco-2 cells. These polarized cells form a tight monolayer and therefore can be used to determine the inhibitory effect of a test compound on bidirectional transport (BDT) of a P-gp probe substrate, e.g., digoxin, from basolateral to apical (B to A) and from apical to basolateral (A to B) (Brouwer et al., 2013). However, this assay requires culturing cells for multiple days to form monolayers. Digoxin transcellular flux is not only mediated by P-gp and passive diffusion, but also by endogenous uptake transporter(s) (Taub et al., 2011; Lee et al., 2014). Varying expression of P-gp and endogenous transporters in different cell lines under different assay conditions may be one of primary contributing factors to high inter-laboratory variability. Furthermore, kinetic analysis of intrinsic inhibitory potency of NMEs on transcellular flux of digoxin by P-gp is complex. Thus, apparent IC_{50} values obtained from conventional analysis may not represent true P-gp inhibitory potency (Zamek-Gliszczynski et al., 2013; Jani and Krajcsi, 2014; Volpe et al., 2014). Alternatively, a vesicular uptake assay is a simpler non-cell-based assay to study P-gp inhibition. The inhibitory effect of a test compound on ATP-dependent uptake of P-gp probe substrate, e.g. N-methyl-quinidine (NMQ), can be measured in MDR1 vesicles (Heredi-Szabo et al., 2013). Unlike the BDT assay, membrane vesicles can be

stored in large quantitate to ensure consistent transporter expression/activity. As inside-out vesicles have direct access to P-gp binding sites, inhibition kinetics follows enzymatic principles. However, vesicular P-gp inhibition can only use the probe substrates with low permeability, e.g., NMQ. This can be a concern when extrapolating inhibition data to digoxin or other P-gp substrates, as P-gp has multiple binding sites (Lee et al., 2010; Sziraki et al., 2011). Currently, calibration of P-gp vesicular inhibition assay, their predictive performance, and inter-laboratory variability is still limited (Ellens et al., 2013; Heredi-Szabo et al., 2013; Fekete et al., 2015).

In this study, we systemically evaluated the predictive performance of two in vitro P-gp inhibition models: vesicular uptake in MDR1 vesicles and BDT in LLC-MDR1 cells using a total of 48 drugs, and digoxin DDI data from 70 clinical studies. A refined cut-off values using ROC analysis were derived for respective in vitro models and compared with the decision frameworks proposed by FDA, EMA and Ellens et al (Ellens et al., 2013). Furthermore, the impact of various IC_{50} calculation methods, and non-specific binding of inhibitor drugs on DDI prediction was evaluated.

Materials and Methods

Chemicals and Reagents

Chemicals

[^{3}H]digoxin (30-40 Ci/mmol) was purchased from Perkin Elmer (Boston, MA). [^{3}H] NMQ (L-000543643-002R001, 73 Ci/mmol) was synthesized by the Labeled Compound Synthesis Department, Merck & Co., Inc., Kenilworth, NJ, USA. Sixty compounds evaluated in P-gp inhibition assays (48 compounds for training set and 12 compounds for test set) were obtained from Sigma (St. Louis, MO), Selleck Chemicals (Pittsburgh, PA), or Cayman Chemical (Ann Arbor, MI). All other reagents were commercially obtained at the highest analytical purity grade.

Cells and membrane vesicles

LLC-PK1 cells and LLC-PK1 cells stably expressing human MDR1 P-gp (LLC-MDR1 cells) were obtained from BD Gentest (Woburn, MA). LLC-MDR1 and LLC-PK1 cells obtained from Netherlands Cancer Institute (Amsterdam, Netherlands) were also used to measure in vitro IC_{50} values for several compounds in the training and test sets. Based on our internal validation, IC_{50} values measured using these two orthogonal cell lines show good correlation (data not shown). The cells were cultured in medium 199 supplemented with 10% fetal bovine serum, 2 mM L-glutamine, 50 U/ml penicillin and 50 μg/ml streptomycin. All cells were maintained at 37°C in an atmosphere of 95% air, 5% CO_2 and 90% relative humidity. Membrane vesicle (lot EUD8G26 and IKATG03) isolated from baculovirus infected *Spodoptera frugiperda* (*Sf*9) cells containing MDR1 P-gp were obtained from Thermo Fisher Scientific (Waltham, MA). Similar time- and ATP-dependent uptake of [^{3}H] NMQ (0.1 μM) were observed in these two lots of

vesicles (data not shown). The sideness of the vesicles were not measured. As nitrogen cavitation method was used to prepare membrane vesicles, we assumed that the vesicles consist of the mixture of equal portion of inside-out and right-side out vesicles (Saito et al., 2009).

In vitro P-gp inhibition assays

BDT inhibition assay in LLC-MDR1 and LLC-PK1 cells

The effects of test compounds on human MDR1 P-gp-mediated efflux transport of digoxin were evaluated using the BDT studies in the LLC-MDR1 and LLC-PK1 cell lines as previously described (Chan et al., 2019). Briefly, cells were cultured in 96-well multiwell insert plates (Millicell®-96, Millipore, Billerica, MA) at 85,000 cells/well and cultured for four days before the study. The compound was tested at seven concentrations in LLC-MDR1 cells as indicated. Cyclosporin A (CsA) (10 μM) was used as a positive control inhibitor. The test compounds or positive control inhibitor at the concentrations indicated were added into both apical (A) and basolateral (B) sides of cell monolayers. Transport buffer was Hanks Balanced Salt Solution (HBSS) with 10 mM (N-[2-Hydroxyethyl] piperazine-N'-[2-ethanesulfonic acid (HEPES), pH 7.4. Transport of digoxin was measured in both absorptive and secretory directions. For absorptive (A to B) transport, the donor dosing solution was added to the apical compartment and for secretory (B to A) transport, donor dosing solution was added to the basolateral compartment. Receiver solution was prepared by adding aliquots of the stock solution of test compound or positive control inhibitor to transport buffer with a final organic solvent concentration of ≤ 1%. Donor dosing solution was prepared by diluting aliquots of radiolabeled and non-radiolabeled [^{3}H]digoxin (final concentration 0.1 μM), and, if applicable, aliquots of test compound or positive control inhibitor stock solutions, into transport buffer at designated

concentrations with a final organic solvent concentration of ≤ 1%. BDT of digoxin without inhibitor was tested in both LLC-MDR1 and LLC-PK1 cells to confirm P-gp-mediated digoxin efflux transport and a sufficient signal-to-noise ratio (efflux ratios (ER, B to A / A to B of apparent passive permeability (P_{app})) in LLC-MDR1 cells/ ER in LLC-PK1 cells ≥ 3). Prior to the transport experiment, cells were washed three times with transport buffer. Donor dosing solution (150 µL) was added to either the apical or basolateral compartment, with receiver solution (150 µL) added to the opposite compartment. At 90 minutes, (50 µL) samples were taken from both sides and scintillant (Ultima Gold, Perkin Elmer, Boston, MA) was added. Radioactivity was determined by liquid scintillation counting in a 2450 MicroBeta2 counter (Perkin Elmer, Waltham, MA). Dextran Texas Red or Lucifer Yellow was used as the markers to test the monolayer integrity. At the end of the incubation, if Dextran Texas Red or Lucifer Yellow in the receiver well was > 6% or >2% of the total concentration, respectively, data were excluded due to poor monolayer integrity**.** The experiments were performed in triplicate.

Vesicular transport inhibition assay in MDR1 P-gp containing membrane vesicles

The inhibitory effect of test compounds on ATP-dependent [^{3}H]NMQ (0.1 µM) uptake was measured in membrane vesicles containing human MDR1 P-gp. The positive control inhibitor (CsA 10 µM) was tested in each assay to confirm the functionality of MDR1 P-gp. Briefly, 19 µl of [^{3}H]NMQ, dissolved in transport buffer (0.25 M sucrose, 10 mM $MgCl_2$, 10 mM Tris-HCl buffer, pH 7.4), were added to 10 µl of MDR1 P-gp containing vesicles (2.5 mg/ml) in a 96-well glass coated plate (Analytical Sales & Services, Flanders, NJ). Then, 1uL of various concentrations of test compounds or CsA (10 µM) were added to each corresponding well. The

incubation plate was preincubated in a water bath for 3 min at 37°C. Uptake was initiated by the addition of 20 µl ATP or AMP containing solution (final concentration of 5 mM ATP or AMP, 10 mM creatine phosphate, and 100 µg/ml creatine phosphokinase in transport buffer), followed by incubation at 37°C for 5 min. Uptake was stopped by the addition of 200 µl ice-cold stop buffer (0.25 M sucrose, 0.1 M NaCl, 10 mM Tris-HCl buffer, pH 7.4), followed by transfer of the reaction mixture to a prewetted 96-well glass fiber filter plate (1.0 µm) (Millipore, Billerica, MA) and application of vacuum. Filters containing the membrane vesicles were washed with 200 µl ice-cold stop buffer five times. The filter plate was dried at room temperature, and 100 µl scintillation fluid (OptiPhase HiSafe3; PerkinElmer, Boston, MA) was added. Radioactivity was determined by liquid scintillation counting. The experiments were performed in triplicate.

Quantification of test compounds via LC/MS/MS analysis

All analyses were performed on a Sciex 4500 triple quadrople mass spectrometer (Toronto, ON, Canada) equipped with electrospray source using multiple reaction monitoring and controlled by Analyst 1.6.2 software. The sample was loaded onto an Acquity UPLC C18 HSS T3, 1.8 µM, 2.1 x 30 mm, 1.8 micron column (Waters, Milford, MA) by means of a Thermo Scientific LX-2 System (Leap Autosampler with Dionex Ultimate 3000 RS Pumps) autosampler controlled by Aria 1.7 software. Chromatography was performed using water, 0.1% formic acid as mobile phase A and acetonitrile, 0.1% formic acid as mobile phase B. The fast gradient profile was 0 to 15 seconds 5% (v/v) B at 0.8 mL/min; 30 seconds ramp to 95% (v/v) B at 1.0 mL/min; 30 seconds to 5% (v/v) B at 0.8 mL/min. Polarity was selected for optimum sensitivity and detection by tandem mass spectrometry was based on precursor ion transitions to the strongest intensity product ions. Samples, standards, and controls were processed with labetalol as the

internal standards. Typical standard curve range was from 7.82 to 2000 nM. LC/MS/MS parameters was shown in **Supplemental Table 1**. Data processing was done using IndigoBio Ascent (Indigo BioAutomation, Carmel, IN).

Clinical DDI data with digoxin and other P-gp probe substrates

Clinical data for 70 DDI studies using digoxin as a probe drug (**Table 1**, training set) and 18 and 6 studies using DE and fexofenadine as probe drugs, respectively (**Table 2**) were collected from the University of Washington DDI database (https://www.druginteractionsolutions.org). A diverse compound set were categorized according to the Provisional Biopharmaceutical Classification System (BCS) obtained from the literature (Wu and Benet, 2005; Benet, 2013; Papich and Martinez, 2015), University of Washington DDI database, and Pharmapendium (https://www.pharmapendium.com). The magnitude of clinical DDIs expressed as AUCR, the ratio of area under the curve (AUC) or $C_{max}R$, the ratio of maximum concentration in plasma (C_{max}) (pharmacokinetic (PK) ratios) with and without inhibitor drugs were collated. A PK ratio of a probe drug ≥1.25 was considered a positive clinical DDI, while a PK ratio <1.25 was defined as a negative outcome. I_1, the mean steady-state total (free and bound) C_{max}, and the fraction unbound in plasma (f_u) for test compounds were collected from the University of Washington DDI database. Extrapolated I_1 values were used, assuming the linear PK, if the values at indicated inhibitor doses were not reported. I_{1u} is the unbound I_1; f_u of 0.01 was used in DDI risk assessment as the worst-case scenario if reported f_u <0.01. I_2, the concentration of drug in the gastrointestinal tract at the given inhibitor dose dissolved in 250 ml was calculated. In addition, a test set of 12 compounds was assembled to evaluate performance of DDI

prediction, and related clinical DDI data of the compounds in test set with orally administered digoxin was collected as indicated above and shown in **Supplemental Table 2**.

Data analysis

IC_{50} calculation

BDT inhibition assay

IC_{50} values for inhibiting digoxin BDT in LLC-MDR1 cells were obtained by fitting the data to equation 1 by nonlinear regression analysis using Graphpad prism (San Diego, CA).

$$\% \text{ Control} = 100 / (1 + I^s / IC_{50}{}^s) \quad \text{Eq. (1)}$$

Where I is the inhibitor concentration (μM), and s is the Hill slope.

Percent control was calculated according to equation 2:

$$\% \text{ Control} = (T_I / T_o) \times 100 \quad \text{Eq. (2)}$$

Where T_I represents net transport of digoxin measured in the presence of various concentrations of the inhibitor; T_o represents the net transport of digoxin in the absence of the inhibitor.

Three methods were commonly used to calculate the net transport of digoxin in LLC-MDR1 cells: 1) net secretory flux (NSF), 2) efflux ratio (ER), and 3) unidirectional flux (UDF) (Balimane et al., 2008; Cook et al., 2010; O'Connor et al., 2015). The NSF evaluates net digoxin transport activity in both absorptive and secretory directions. The ER describes the ratio of P_{app} in secretory (P_{app} B to A) over the absorptive direction (P_{app} A to B). The UDF (B to A) describes transport activity of digoxin in the secretory direction (basolateral to apical: B to A).

NSF in LLC-MDR1 cells was calculated according to equation 3:

$$NSF = (\%Transport\ B\ to\ A) - (\%Transport\ A\ to\ B) \quad \text{Eq. (3)}$$

%Transport was calculated by dividing the amount of digoxin measured in the receiver compartment by the sum of digoxin measured in both receiver and donor compartments.

ER in LLC-MDR1 cells was calculated via Eq. (4)

$$ER = P_{app}\ (B\ to\ A)/\ P_{app}\ (A\ to\ B) \quad \text{Eq. (4)}$$

P_{app} is calculated by Eq (5):

$$P_{app} = (\text{Volume of receiver chamber}) / (A \times C_o) \times \Delta\ \text{Conc} / \Delta t \quad \text{Eq. (5)}$$

The volume of receiver chamber is 0.15mL; transwell membrane area (A) is 0.11cm^2; C_o is the sum of the probe substrate concentration (digoxin) measured in the donor plus receiver compartments at 1.5 hr; Δ in concentration is the concentration in the receiver compartment at 1.5 hr; and Δ in Time is the incubation time (1.5 x 60 x 60 = 5400s). P_{app} is expressed as 10^{-6} cm/s.

UDF (B to A) in LLC-MDR1 cells was calculated by Eq. (6)

$$UDF\ (B\ to\ A) = (P_{app}\ BA_i - P_{app}\ BA_p)/ (P_{app}\ BA_o - P_{app}\ BA_p) \quad \text{Eq. (6)}$$

Where $P_{app}\ BA_i$ represents P_{app} receiver B to A with inhibitor; $P_{app}BA_o$, B to A receiver without inhibitor, $P_{app}\ BA_p$, B to A receiver with positive control inhibitor (CsA, 10 μM).

Considering potential non-specific binding of inhibitor compounds to assay plates and cells, a separate set of IC_{50} analysis was performed by correcting IC_{50} values obtained from the nominal inhibitor concentrations based on the recovery of inhibitor drugs measured in the incubation.

The recovery of the inhibitor drugs (R_i) was measured in well via LC/MS/MS analysis under the BDT assay conditions as described above. R_i was calculated by dividing the sum of the inhibitor concentration measured in the donor (C_{donor}) and receiver ($C_{receiver}$) compartments at the end of the incubation divided by C_{dose}, the concentration of test compound in dosing solution at time zero. The concentrations of the inhibitor drugs remained in the cells were not measured.

$$R_i = [(C_{donor} + C_{receiver}) / C_{dose}] \quad \text{Eq. (7)}$$

As the recovery of the test compound was calculated only based on the concentration of the test compound in apical and basolateral chambers, lower recovery suggested possible non-specific binding of the test compound to cell monolayers and /or the assay plates. If R_i was < 0.7, in vitro IC_{50} values measured based on nominal inhibitor concentrations were corrected based on Eqs (8-10), respectively. If R_i was ≥ 0.7, the IC_{50} values were not corrected.

$$IC_{50}(NSF_c) = IC_{50}(NSF) \times R_i \quad \text{Eq. (8)}$$

$$IC_{50}(ER_c) = IC_{50}(ER) \times R_i \quad \text{Eq. (9)}$$

$$IC_{50}(UDF_c) = IC_{50}(UDF) \times R_i \quad \text{Eq. (10)}$$

Where $IC_{50}(NSF)$, $IC_{50}(ER)$, and $IC_{50}(UDF)$ represent IC_{50} values estimated by calculating digoxin net transport using NSF, ER, and UDF, respectively. $IC_{50}(NSF_c)$, $IC_{50}(ER_c)$ and $IC_{50}(UDF_c)$, represent corrected IC_{50} values based on the recovery of test compounds.

Vesicular Inhibition Assay

MDR1 P-gp-mediated NMQ transport was estimated based on ATP-dependent [^{3}H] NMQ uptake in MDR1 vesicles calculated by subtracting uptake of [^{3}H] NMQ in presence of AMP

from uptake in presence of ATP. Percent control in vesicular inhibition assay is calculated according to equation 11:

$$\% \text{ Control} = (V_I / V_o) \times 100 \qquad \text{Eq. (11)}$$

where V_I represents ATP-dependent uptake rate of NMQ measured in the presence of various concentrations of the test compound; V_o represents ATP-dependent uptake rate in the absence of the test compound. As [^{3}H] NMQ exhibited minimal active uptake in control vesicles and in MDR1 vesicles in the presence of ATP and positive control inhibitor (10 µM CsA) (**Supplemental Figure S1**), no control vesicles or positive control inhibitor were used to correct P-gp-mediated NMQ transport rate.

The IC_{50} for inhibition of MDR1 P-gp-mediated NMQ vesicular transport, $IC_{50}(V)$, was obtained by fitting the data to equation (1).

Inhibitor drugs may have non-specific binding to the assay plates and vesicles. As our vesicular inhibition assay was conducted in glass-coated plates and the surface area as well as the volume of vesicles are small (Nervi et al., 2010) at the given vesicle amount (25 µg) in this assay, non-specific binding of inhibitor drugs may have negligible impact on inhibitor concentrations in the incubation (nominal extravesicular concentration). This was confirmed by our pilot studies for several highly bound inhibitor drugs following the measurement of the binding of inhibitor drugs to vesicles using ultracentrifugation method (unpublished observations). Therefore, in this study, IC_{50} estimated based on nominal inhibitor concentrations was used for vesicular inhibition assay without the correction for non-specific binding.

Statistical methods and the criteria to define optimized cut-off values

A binary classification analysis was conducted on clinical digoxin DDI studies for the compounds in the training set (**Table 1**) to derive the optimal cut-off values to predict digoxin DDIs based on three static models: 1) I_2/IC_{50}(FDA, 2020); 2) I_1/IC_{50} or I_2/IC_{50}(Agarwal et al., 2013); and 3) I_{1u}/IC_{50}, or I_2/IC_{50}(EMA, 2012).

IC_{50} values measured in the BDT and vesicular inhibition assays and calculated using various methods (IC_{50}(NSF), IC_{50}(NSF_c), IC_{50}(ER), IC_{50}(ER_c), IC_{50}(UDF), IC_{50}(UDF_c) and IC_{50}(V)) were used in this analysis, and the predictive performances were compared.

The possible outcomes were: 1) true positive (TP), in vitro data predicts a positive digoxin DDI and the prediction is in agreement with a positive clinical DDI (AUCR $\geq$1.25 or $C_{max}R \geq 1.25$); 2) false negative (FN), in vitro data predict a negative digoxin DDI, but the prediction is not in agreement with positive clinical data; 3) true negative (TN), in vitro data predict a negative DDI and the prediction is in agreement with a negative clinical DDI (AUCR <1.25 and $C_{max}R < 1.25$); and 4) false positive (FP), in vitro data predict a positive DDI, but it is not observed in clinic.

Performance metrics used in the analyses are defined and calculated as below:

Sensitivity = True positive rate =TP/P Eq(12)

Specificity = True negative rate =TN/N Eq(13)

Overall accuracy = (TP +TN)/S Eq(14)

Average accuracy=(Sensitivity+Specificity)/2 Eq(15)

False negative rate =FN/P Eq(16)

False positive rate = 1-Specificity Eq(17)

S is the total number of digoxin DDI studies, P the number of positive clinical DDI studies (AUCR ≥1.25 or $C_{max}R \geq 1.25$), and N is the number of negative studies (AUCR <1.25 and $C_{max}R < 1.25$).

The Receiver Operating Characteristics (ROC) curve, which is a plot with the sensitivity versus false positive rate (1-specificity) at various cutoffs, is used to check and visualize the overall performance of a binary classifier. The AUC of a ROC curve (AUC_{ROC}) is one of the most important evaluation metrics. An AUC_{ROC} =1 indicates a perfect separation of the two classes (positive or negative), and an AUC_{ROC} =0.5 means that the classifier is the same as randomly splitting the outcomes. We first apply ROC analysis on the one-dimensional classifiers: the static models of I_2/IC_{50}(model1). The ROC curve is generalized to the two-dimensional classifiers for static model I_1/IC_{50} or I_2/IC_{50}(model 2) and I_{1u}/IC_{50} or I_2/IC_{50}(model 3), respectively, by plotting the sensitivity versus the sorted false positive rate after pooling all the possible two-dimensional cutoffs.

In the analysis, the AUC_{ROC}s (Hanley and McNeil, 1982; Robin et al., 2011) and their 95% confidence intervals (Hanley and McNeil, 1982) were used to evaluate and compare the overall classification power of proposed classifiers/tests. The classifier with best AUC_{ROC} was selected. The computation was conducted using open-source programming language **R** (http://www.R-project.org).

A cutoff is one or a set of values that defines a positive or negative outcome in a binary classifier. Therefore, the cutoff also defines the performance metrics. ROC analysis uses these metrics to determine the optimal cutoff. For a chosen classifier, the following method was used to find the optimal cutoffs (*C*):

$$C = argmax_{c \in \{\theta : Sensitivity(\theta) \geq 0.75\}} Specificity\ (c) \qquad \text{Eq(18)}$$

Where "*argmax*" is an operation that finds the argument that gives the maximum value from a target function. $Sensitivity(c)$ and $Specificity\ (c)$ are defined for a given cutoff *c*. This method maximizes *Specificity* by searching in the cutoff space, where a *Sensitivity* of ≥ 0.75 is guaranteed.

The predictive performance using derived cut-offs based on ROC analysis was compared with the following static models and recommended cut-off values by regulatory agencies and Ellens et al (Ellens et al., 2013): 1) $I_2/IC_{50} \geq 10$ (FDA, 2020); 2) $I_2/IC_{50} \geq 45$ (Ellens et al., 2013); 3) $I_1/IC_{50} \geq 0.1$ or $I_2/IC_{50} \geq 10$ (Agarwal et al., 2013); 4) $I_1/IC_{50} \geq 0.03$ or $I_2/IC_{50} \geq 45$ (Ellens et al., 2013); and 5) $I_{1u}/IC_{50} \geq 0.02$ or $I_2/IC_{50} \geq 10$ (EMA, 2012). The cut-offs recommended by FDA and EMA were empirical values, while the cut-off proposed by Ellens et al. was derived based on ROC analysis of P-gp IC_{50} data for 15 compounds generated by 23 laboratories using four in vitro systems: Caco-2 cells, LLCPK1-MDR1, MDCKII-MDR1, and MDR1 vesicles. P-gp probe substrates were digoxin for polarized cell-lines and NMQ or vinblastine for MDR1 vesicles (Ellens et al., 2013). After optimizing the cut-off values for different models using the data from the training set, the performance of selected models and optimized cut-offs were further verified on a test set with 12 compounds not included in the training set.

Results

Comparison of in vitro P-gp IC_{50} values measured in the BDT and vesicular inhibition assays

In vitro IC_{50} values of 48 compounds in the training set measured in BDT and vesicular inhibition assays are summarized in **Table 3**, and IC_{50} plots of all test compounds measured in BDT assay using NSF, ER and MDR1 vesicular inhibition assay are shown in **Supplemental Figure S2-1 and S2-2**. For the BDT assay, IC_{50} values were calculated using NSF, ER, and UDF, with and without correction for non-specific binding. Obtained IC_{50} data (IC_{50}(NSF), IC_{50}(NSF_c), IC_{50}(ER), IC_{50}(ER_c), IC_{50}(UDF), and IC_{50}(UDF_c)) were compared with IC_{50} values from vesicular inhibition assay (IC_{50}(V)), respectively (**Figure 1**). In cases where the IC_{50} values could not be determined due to the lack of more than 50% inhibition (not an inhibitor at highest tested concentration or the solubility limit), the highest inhibitor concentrations tested were used as surrogate IC_{50} values for the purpose of comparison across different assays and methods.

The variability of vesicular vs. BDT inhibition assay was evaluated. **Supplemental Table 3** compared the variability the IC_{50} values for verapamil and quinidine conducted in three independent studies in both MDR1 vesicles and LLC-MDR1 cells (IC_{50}(V) vs. IC_{50}(NSF)). Based on this limited dataset, IC_{50} values measured in MDR1 vesicles were less variable (lower CV%) than in BDT assay. This is also supported by the data collected in the literature when the same assay systems and probe substrates were used. we have also compared IC_{50}(V) values generated in this study with those reported in the literature using MDR1 vesicles. The results were summarized in **Supplementary Table 4**. The IC_{50} values reported in literature using MDR1 vesicles are still limited (25 out of 48 compounds have IC_{50} data reported in the

literature). Nevertheless, 6 out of 25 compounds showed >10-fold difference (10.8-22-fold) on IC_{50}(V)s compared to those reported values. Despite that, such inter-laboratory variability is lower than reported by others in BDT assays using different P-gp transfected cells and Caco-2 cells (Bentz et al., 2013). Further studies are needed to confirm this observation.

To compare IC_{50}(NSF) with IC_{50}(V) (**Figure 1A-1**), the IC_{50}(NSF) values for 14 out of 48 compounds were 5-fold higher than the IC_{50}(V), in which 10 compounds had at least 10-fold higher IC_{50}(NSF) than IC_{50}(V) ranging from 10.2-fold for velpatasvir to 224-fold for fidaxomicin. In contrast, IC_{50}(V) values for only 3 compounds were 5-fold higher than IC_{50}(NSF), ranging from 7.2-fold for felodipine to 10-fold for alogliptin. The difference for alogliptin was attributed to the difference of the highest concentrations tested, as no inhibition was observed in both assays at the highest concentrations tested. As shown in **Figure 1C-1**, IC_{50}(ER) for 12 compounds were more than 5-fold lower than IC_{50}(V) (ranging from 6- to 24-fold), whereas IC_{50}(ER) values of 9 compounds were at least 5-fold higher than IC_{50}(V) (ranging from 5 to 224-fold (fidaxomicin)). IC_{50}(UDF) values of 15 compounds were at least 5-fold higher than IC_{50}(V), while only 3 compounds had 5-fold lower IC_{50}(UDF) than IC_{50}(V) (**Figure 1 E-1**). Overall, there is a poor correlation between IC_{50}(NSF) and IC_{50}(V) (**Figure 1 A-2**), IC_{50}(ER) and IC_{50}(V) (**Figure 1 C-2**), as well as IC_{50}(UDF) and IC_{50}(V) (**Figure 1 E-2**). The correlation was improved, when alogliptin and maraviroc were excluded from the analysis (**Figure 1 A-3**, **C-3**, **E-3**).

In the BDT assays, the IC_{50} values obtained after correction for non-specific binding of inhibitor compounds ($IC_{50}(NSF_c)$, $IC_{50}(ER_c)$, and $IC_{50}(UDF_c)$) were compared with IC_{50}(V), respectively, and the results are shown in **Table 3** and **Figure 1B-1**, **1D-1**, and **1F-1**. As described in the

method section, non-specific binding of inhibitor drugs was not corrected for IC_{50}(V). The IC_{50} values corrected for non-specific binding were generally lower for the compounds which had high non-specific binding to the assay plates and /or cells as evidenced by low recovery in the incubation. The stability of test compound during incubation was confirmed (data not shown). $IC_{50}(NSF_c)$ for 15 compounds were 5-fold lower than IC_{50}(V). On the contrary, $IC_{50}(NSF_c)$ values for 10 compounds were 5-fold higher than IC_{50}(V), in which 6 compounds had 10-fold higher $IC_{50}(NSF_c)$ than IC_{50}(V) (**Table 3**, **Figure 1B-1**). Compared to IC_{50}(ER), $IC_{50}(ER_c)$ values trended towards lower than IC_{50}(V): $IC_{50}(ER_c)$ of 23 compounds were more than 5-fold lower than IC_{50}(V), where 18 out of 23 compounds showed at least 10-fold lower $IC_{50}(ER_c)$ than IC_{50}(V). In contrast, only 7 compounds showed 5-fold higher $IC_{50}(ER_c)$ than IC_{50}(V). Similarly, $IC_{50}(UDF_c)$ for 15 compounds were more than 5-fold lower than IC_{50}(V), whereas $IC_{50}(UDF_c)$ of 10 compounds were more than 5-fold higher than IC_{50}(V). Overall, the correction for non-specific binding did not improve, but rather reduced the correlation between the IC_{50} values measured by the BDT and vesicular inhibition assays. There was no correlation between $IC_{50}(NSF_c)$ and IC_{50}(V) (**Figure 1**, **B-2**; R^2=0.027, P>0.05); $IC_{50}(ER_c)$ and IC_{50}(V) (**Figure 1**, **D-2**; R^2=0.032, P>0.05), as well as $IC_{50}(UDF_c)$ and IC_{50}(V) (**Figure 1 F-2**, R^2=0.053, P>0.05). Excluding alogliptin, maraviroc, and captopril resulted in an improved, but yet poor correlation (**Figure 1**, **B-3**; R^2=0.277, P<0.001; **D-3**, R^2=0.299, P<0.001; **F-3**, R^2=0.324, P<0.001). Among the 48 compounds tested, the difference in observed IC_{50} values using the BDT and vesicular assays for BCS I and III compounds appeared to be less than those for BCS II and IV compounds, despite the lack of a clear trend and a limited dataset.

Comparison of IC_{50} values in the BDT assays using different IC_{50} calculation methods

Figure 2 showed IC_{50} values measured in the BDT assays calculated by NSF, ER, and UDF. IC_{50}(NSF) of all test compounds were comparable to those of IC_{50}(UDF) with the difference less than 3- to 5-fold, except for one outlier, vemurafenib (BCS IV compound): whose IC_{50}(NSF) was 35-fold higher than IC_{50}(UDF) (**Figure 2A-1**, **Table 3**). IC_{50}(ER) exhibited a trend of lower than IC_{50}(NSF) (**Figure 2B-1**, **Table 3**): IC_{50}(ER) values of 11 compounds were more than 5-fold lower than IC_{50}(NSF), in which 6 compounds were more than 10-fold lower than those of IC_{50}(NSF) (**Figure 2B-1**, **Table 3**). Likewise, IC_{50}(ER) values also trended towards lower than IC_{50}(UDF), except vemurafenib's IC_{50}(ER), which was 20-fold higher than IC_{50}(UDF). IC_{50}(UDF) of 12 compounds were more than 5-fold higher than IC_{50}(ER), where 10 out of 12 compounds had more than 10-fold higher IC_{50}(UDF) (**Figure 2C-1**). In general, there is a good correlation when IC_{50}(NSF) vs. IC_{50}(UDF), IC_{50}(NSF) vs. IC_{50}(ER), and IC_{50}(UDF) vs. IC_{50}(ER) were compared. Excluding captopril and maraviroc, two compounds with IC_{50} values (IC_{50}(NSF), IC_{50}(ER), IC_{50}(UDF)) greater than 1000 μM, there was still a good correlation between IC_{50}(NSF) and IC_{50}(ER) (R^2=0.875, P<0.0001), followed by IC_{50}(NSF) vs. IC_{50}(UDF) (R^2= 0.676, P<0.001), and IC_{50}(UDF) vs. IC_{50}(ER) (R^2=0.513, P<0.0001) (**Figure 2**, **A-3**, **B-3**, and **C-3**).

Clinical Digoxin DDI data for ROC analysis training and test sets

Table 1 summarized 70 clinical digoxin DDI studies, I_1, I_2, f_u, AUCR and $C_{max}R$ for 48 compounds with measured in vitro P-gp IC_{50} values (**Table 3**). These data were used as the training set for ROC analysis. Only clinical DDI data with orally administered digoxin were collected. For certain compounds, multiple clinical DDI datasets at different perpetrator dose regimen and study design were included. In this dataset, 48 clinical DDIs were positive (AUCR and/or $C_{max}R \geq 1.25$) and 22 DDI data were negative (AUCR and $C_{max}R$ <1.25). For 48

compounds tested, 6, 23, 6, and 9 compounds were classified as BCS Class I, II, III, and IV, respectively. Due to inconsistent information from BCS classification database and literature, 2 compounds were classified as BCS Class I or II, and another 2 compounds were classified as BCS Class II or IV. We also conducted in vitro P-gp inhibition studies for additional 12 compounds as the test set. In vitro IC_{50} values obtained (IC_{50}(V), IC_{50}(ER), IC_{50}(NSF)) were shown in **Supplemental Table 5** and related clinical digoxin DDI data (8 positive and 4 negative) are summarized in **Supplemental Table 2**. These data were used as the test set to assess the predictive performance of the models and the cut-off values derived by ROC analysis.

Comparison of ROC analysis using various static P-gp DDI prediction models and IC_{50} measurement

To select the optimal model and IC_{50} assay to predict digoxin DDIs, ROC analysis was conducted with three P-gp DDI prediction models: 1) I_2/IC_{50}; 2) I_1/IC_{50} or I_2/IC_{50}; and 3) I_{1u}/IC_{50}, or I_2/IC_{50} using IC_{50} values (IC_{50}(V), IC_{50}(NSF), IC_{50}(NSF_c), IC_{50}(ER), IC_{50}(ER_c), IC_{50}(UDF), and IC_{50}(UDF_c)) obtained from various assays and calculation methods. AUC_{ROC} values and respective ROC curves obtained from this analysis were shown in **Figure 3** and **Figure 4**, respectively. Despite the lack of statistical significance (possibly due to the relatively small sample sizes), model 1 (I_2/IC_{50}) with IC_{50}(V) demonstrated a trend of highest AUC_{ROC}, followed by IC_{50}(ER), IC_{50}(NSF), and IC_{50}(UDF). Similar observations were also found in model 2 (I_1/IC_{50}, or I_2/IC_{50}), and model 3 (I_{1u}/IC_{50}, or I_2/IC_{50}) (**Figure 3**). In all three models, IC_{50} measurement in the BDT assay after the correction of non-specific binding of inhibitor drugs (IC_{50}(ER_c), IC_{50}(NSF_c), and IC_{50}(UDF_c)) showed a trend of lower AUC_{ROC} than respective IC_{50} measurement without the correction of nonspecific binding ((IC_{50}(ER), IC_{50}(NSF), and

IC_{50}(UDF)). As shown in **Figure 4**, the ROC curve was well-defined for one dimensional binary classifier (model 1: I_2/IC_{50}), but not for two dimensional classifiers (models 2 and 3: I_1/IC_{50} or I_2/IC_{50}; I_{1u}/IC_{50} or I_2/IC_{50}), as their true positive rate and true negative rate were not monotonically related.

Based on ROC analysis, optimal discrimination thresholds that achieve the highest specificity constrained to a minimal sensitivity of 0.75 was obtained for these models, and their predictive performance were summarized in **Table 4**. Only IC_{50}(V), IC_{50}(ER), and IC_{50}(NSF) values were used in this analysis, as they had relatively higher AUC_{ROC} values. The predictive performance of each models with optimal thresholds derived from the ROC analysis was compared with respective models and the cut-off values recommended by regulatory agencies (EMA, 2012; Agarwal et al., 2013; FDA, 2020) and Ellens et al (Ellens et al., 2013) (**Table 4**).

In model 1 (I_2/IC_{50}), the cut-off value for I_2/IC_{50} using IC_{50}(V) was 25.9 with the percentage of TP, TN, FP, and FN of 75%, 59.1%, 40.9%, and 25%, respectively, and the average and overall accuracy of 0.67 and 0.7 (**Table 4-1**). Using the same data set, the cut-off values recommended by FDA ($I_2/IC_{50} \geq 10$) resulted in lower FN% (16.7%) but higher FP% (54.5%), while the cut-off by Ellens et al ($I_2/IC_{50} \geq 45$) showed higher FN% (31.3%). Using IC_{50}(ER), the obtained cut-off value of 13.7 yielded the percentage of TP, TN, FP, and FN of 83.3%, 50%, 50%, and 16.7%, respectively, and the average and overall accuracy of 0.667 and 0.729, respectively. These results were generally comparable to those using FDA cut-off value ($I_2/IC_{50} \geq 10$), while the cut-off value by Ellens et al resulted in a higher FN% (35.4%), despite a relatively lower FP% (36.4%). Likewise, the cut-off value derived based on IC_{50}(NSF) ($I_2/IC_{50} \geq 9.3$) is comparable to the one from FDA ($I_2/IC_{50} \geq 10$) with similar accuracy, and the percentage of TP, TN, FP, and

FN, while the cut-off value by Ellens et al ($I_2/IC_{50} \geq 45$) resulted in lower accuracy, and higher FN%.

In model 2 (I_1/IC_{50}, or I_2/IC_{50}), the cut-off values obtained with IC_{50}(V), IC_{50}(ER), and IC_{50}(NSF) was (0.032, 40), (0.081, 26.7), and (0.026, 10), respectively (**Table 4-2**). The accuracy with IC_{50}(V) and IC_{50}(NSF) was the same, while the accuracy of IC_{50}(ER) was lower with higher FP% (50%). Using the cut-off value of $I_1/IC_{50} \geq 0.1$, or $I_2/IC_{50} \geq 10$, all three IC_{50} dataset had similar accuracy with lower FN% for IC_{50}(V) and IC_{50}(ER). Using the cut-off value of $I_1/IC_{50} \geq 0.03$, or $I_2/IC_{50} \geq 45$ by Ellens et al, IC_{50}(V) had the same FP% and FN% as the cut-off derived from ROC analysis, but IC_{50}(ER) and IC_{50}(NSF) resulted in lower accuracy and higher FN%. In addition, we have further compared the cut-off values and predictive performance derived from our ROC analysis using IC_{50}(V) data with those reported by Ellens et al.(Ellens et al., 2013) using only MDR1 vesicular IC_{50} dataset for 15 compounds generated in five laboratories using either NMQ or vinblastine as in vitro probes (data shown in Supplemental Table 2 of Ellens et al.). In addition, the cut-off and predictive performance using vesicular IC_{50} data generated in a single lab (lab 20, Supplemental Table 2 of Ellens et al.) was compared. In brief, the cut-off value ($I_1/IC_{50} \geq 0.08$, $I_2/IC_{50} \geq 501$) derived from all MDR1 vesicular IC_{50} data by Ellens et al. resulted in a FN% and FP% of 33% and 8%, respectively (FN% and FP % of 36% and 33%, respectively in a test set). The cut-off value ($I_1/IC_{50} \geq 0.01$, $I_2/IC_{50} \geq 89$) derived from vesicular IC_{50} data in a single lab (lab 20) showed a FN % and FP % of 36% and 18%, respectively. In contrast, the cut-off value derived from our vesicular data ($I_1/IC_{50} \geq 0.032$, $I_2/IC_{50} \geq 40$; **Table 4-2**) showed lower FN% (25%), but higher FP% (36.4%). The difference in the cut-off values derived from different laboratories may be attributed to inter-laboratory variability of IC_{50} data measured with MDR1 vesicles and different set of training compounds

used in the calibration. This highlights the need to calibrate MDR1 vesicular assays for P-gp DDI prediction. The difference in probe substrates (NMQ vs. vinblastine), the source of membrane vesicles (P-gp expression levels, the ratio of inside-out to right-side-out vesicles), and assay conditions may contribute to such inter-laboratory variability.

In model 3 (I_{1u}/IC_{50}, I_2/IC_{50}), the cut-off values derived based on IC_{50}(V), IC_{50}(ER), and IC_{50}(NSF) was (0.00141, 3334), (0.00177, 62), and (0.00052, 94), respectively, which was approximately 12-38 lower in I_{1u}/IC_{50}, but 6-333-fold higher in I_2/IC_{50} than the threshold recommended by EMA ($I_{1u}/IC_{50} \geq 0.02$ or $I_2/IC_{50} \geq 10$) (EMA, 2012). Under the pre-defined sensitivity of 0.75 (FN% 25%), ROC refined cut-off value with IC_{50}(V) resulted in lowest FP% (18.2%) comparing to that of IC_{50}(ER) (45.5%) and IC_{50}(NSF) (31.8%). EMA cut-off value generally resulted in lower FN% (14.6-25%), but higher FP% (36.4-54.5%). The list of FN and FP compounds in this training set under different models and cut-off values was summarized in **Supplemental Table 5**.

A total of 12 compounds were selected as an independent test set to evaluate the performance of ROC refined cut-offs using IC_{50}(V), IC_{50}(ER), and IC_{50}(NSF) (**Supplemental Table 6**) and clinical digoxin (oral dose only) DDI data (**Supplemental Table 2**; 8 positive, 4 negative DDIs). The predictive performance of the test set was summarized in **Supplemental Table 7**. Within this limited data set, for I_2/IC_{50} model, ROC derived cut-off values for both IC_{50}(V), and IC_{50}(NSF) resulted in no FN, while IC_{50}(ER) showed 12.5% FN. The cut-off values by FDA and Ellens et al. provided reasonably good performance but higher FN% for the cut-off by Ellens et al. Similar results were shown for model 2 (I_1/IC_{50} or I_2/IC_{50}): IC_{50}(V) and IC_{50}(NSF) did not show FN, while IC_{50}(ER) had FN% of 12.5. For model 3 (I_{1u}/IC_{50}, I_2/IC_{50}), a higher FN% for

ROC derived cut-offs than the one recommended by EMA was observed for IC_{50}(V), IC_{50}(ER) and IC_{50}(NSF). The list of FN and FP prediction of the compounds in the test set was summarized in **Supplemental Table 8**.

Prediction of P-gp-mediated DDIs for selected compounds using DE and fexofenadine as clinical probe drugs

To understand whether in vitro P-gp IC_{50} values measured using digoxin and NMQ as probes and the cut-off values derived from digoxin clinical DDI data can be used to predict DDIs for other P-gp probe substrates, we extended our analysis by evaluating predictive performance of 23 P-gp related clinical DDI data using DE (17) and fexofenadine (6) as probes (**Table 2**, 19 positive 4 negative DDIs, respectively) for 15 inhibitor drugs with IC_{50} values measured in our studies (**Table 3)**. Only the performance of model 1 (I_2/IC_{50}) was evaluated, as gut P-gp is the major site for P-gp related DDIs with DE and fexofenadine (Chu et al., 2018a; Chu et al., 2018b; Li et al., 2017; Lappin et al., 2010). In this limited dataset, applying the cut-off values derived from ROC analysis based on clinical digoxin DDI data, the FN% of DDI prediction for DE and fexofenadine using IC_{50}(V), IC_{50}(ER), and IC_{50}(NSF) was 15.8%, 10.5%, and 0%, respectively, while the FP% was 50%, 50%, and 75%, respectively (**Table 5**). Likewise, FDA cut-off value resulted in low FN%, but high FP%. The cut-off value by Ellens et al using IC_{50}(V) and IC_{50}(ER) showed FN% of 21.1% and 26.3%, and FP% of 25% and 25%, respectively, whereas IC_{50}(NSF) data resulted in high FN% (52.6%). The list of compounds with FN and FP prediction in this dataset was summarized in **Supplemental Table 9**. In model 1, three FN predictions were observed using IC_{50}(V). However, the magnitude of clinical DDIs ranged only 1.26- to 1.68-fold, indicating weak to moderate DDIs. Two FP predictions using IC_{50}(V) were

atorvastatin and ritonavir, BCS Class II and IV compound, respectively, which is likely caused by lower inhibitor concentration in the gut due to low solubility.

Discussion

To improve the prediction of P-gp related DDIs, we have systemically calibrated our P-gp inhibition assays (LLC-MDR1 cells and MDR1 vesicles). This is for the first time a side-by-side comparison of two P-gp inhibition assays conducted in the same laboratory. To our knowledge, the numbers of compounds calibrated in MDR1 vesicles in this study are larger than those reported in the literature to date (Ellens et al., 2013; Heredi-Szabo et al., 2013; Fekete et al., 2015).

There is a system-dependent difference of IC_{50} values between BDT and vesicular inhibition assays. Despite a limited dataset, it appears that more difference in IC_{50}s was observed for BCS class II and IV compounds. For instance, IC_{50}(NSF), IC_{50}(ER), and IC_{50}(UDF) of velpatasvir and daclatasvir were more than 10- and 20-fold higher than IC_{50}(V) and resulted in false negative DDI predictions. IC_{50}(V) of fidaxomicin was at least 224-fold lower than IC_{50}(NSF), IC_{50}(ER), and IC_{50}(UDF). The correction for nonspecific binding in BDT assay did not reduce such difference. In contrast, a FP prediction was observed for fidaxomicin using IC_{50}(V). The mechanisms for such difference are not known, likely caused by substrate-dependent inhibition (digoxin vs. NMQ). Furthermore, in inside-out membrane vesicles, inhibitor drugs have direct access to P-gp binding sites, while in BDT assay, inhibitors need to permeate across the lipid bilayers to access P-gp binding sites located in the cytosolic leaflet of the plasma membrane. This could yield a difference of apparent IC_{50} values for a poorly permeable inhibitor drug, like fidaxomicin. For BDT assay, there were also some differences in IC_{50} values generated by different calculation methods (ER, NSF, and UDF) even if the same dataset was used. However, such difference was generally less profound than those between BDT and vesicular inhibition

assay. Currently, there is no consensus on the optimal method to calculate P-gp IC_{50} in BDT assay. The model-based approach may provide a more mechanistic and accurate estimation of IC_{50} and K_i values (Kishimoto et al., 2016).

The predictive performances of various static models, IC_{50} measurement and calculation methods were compared. Among 3 static models evaluated, inclusion of I_1/IC_{50}(model 2) or I_{1u}/IC_{50}(model 3) did not provide superior predictive performance than model 1(I_2/IC_{50} only) based on ROC_{AUC}, FN%, FP% and the accuracy of DDI prediction. Furthermore, different from model 1, 2D ROC curves in models 2 and 3 were in zig-zig shape, suggesting that the standard ROC analysis for 2D classifiers was not well-defined and might need further statistical modeling of the 2D predictors to follow suit. This may also be attributed to the fact that only orally administered digoxin clinical DDI data were collected in this analysis. This observation, similar to others (Poirier et al., 2014; Zhou et al., 2019), suggested that inhibition of intestinal, but not renal P-gp is the major DDI mechanism for orally administered digoxin. Therefore, model 1 (I_2/IC_{50}) is sufficient to predict DDIs for orally administered P-gp substrates for a simplistic and practical considerations, whereas models 2 and 3 did not improve prediction accuracy, but rather add uncertainty for highly bound inhibitor drugs when f_u cannot be accurately measured.

Based on $AUC_{ROC,}$ IC_{50}(V) appeared to show the trend of better predictive performance, followed by IC_{50}(ER), IC_{50}(NSF), and IC_{50}(UDF) across all 3 static models evaluated, despite the lack of statistical significance, due to the limited sample sizes. Overall, the optimal discrimination thresholds derived from this training set for respective in vitro IC_{50} data had a minimal sensitivity of 0.75, and highest specificity. For example, in model 1 (I_2/IC_{50}), IC_{50}(V), IC_{50}(ER) and IC_{50}(NSF) had similar accuracy values. IC_{50}(ER) demonstrated lowest FN%, but

highest FP%, while IC_{50}(V) and IC_{50}(NSF) showed comparable FN% and FP%. Compared to the cut-off values recommended by FDA and Ellens et al., the cut-offs derived from our ROC analysis exhibited better or comparable predictive performance, which highlighted the need to calibrate in vitro systems to provide more accurate DDI prediction. The static models and ROC derived cut-off values were further validated with an independent test set of 12 compounds, confirming a good predictive performance.

In training set, there were a total of 12 FN prediction for IC_{50}(V) using optimized I_2/IC_{50} cut off value (25.9). For these FN predictions, 5 out of 12 had AUCR or $C_{max}R$ <1.5, suggesting weak DDIs, while 7 out of 12 had AUCR or $C_{max}R$ ranged 1.5-2. For these 7 studies, the mechanisms of underprediction were not well understood. For captopril, FN prediction was observed in all models and assays using either optimized cut-offs or values recommended by FDA, EMA, and Ellens et al. Carvedilol underprediction was only observed in DDI studies with female subjects. Underprediction of flibanserin DDIs could be attributed to substrate- and/or system- dependent difference of IC_{50} measurement, as FN prediction was not shown using IC_{50}(NSF) and IC_{50}(ER).

As digoxin is neither a specific nor sensitive P-gp probe (Taub et al., 2011; Lee et al., 2014), DE has been recommended by regulatory agencies as an alternative clinical probe for gut P-gp inhibition (EMA, 2012; FDA, 2020). However, DE was neither stable in cell-based assays (Chu et al., 2018a), nor showing robust transport in MDR1 vesicles (unpublished observations) likely due to higher P_{app} and non-specific binding. This has precluded the use of DE as an in vitro probe for P-gp inhibition. In this study, we explored the feasibility of using NMQ and digoxin and the cut-offs derived from digoxin clinical DDI studies (model 1) to predict P-gp related DDIs for DE and fexofenadine, another P-gp substrate (Chu et al., 2018b). Due to limited in vitro and clinical

data, only 23 clinical studies were selected in this analysis. Overall, a lower FN% was observed for IC_{50}(V) and IC_{50}(ER) with higher FP%. As only a total of 4 negative DDI data were available in this dataset, it may not be feasible to accurately assess the predictive performance, especially for specificity. More compounds should be included in future studies.

Beyond high IC_{50} variability, there are several issues and knowledge gaps for P-gp DDI prediction. 1) Prediction of DDIs for perpetrators with low solubility. For class II /IV compounds, the solubility of inhibitor drugs can be much lower than I_2. I_2/IC_{50} model may not be feasible to derisk DDIs due to the inability to test the inhibitor concentration beyond the solubility limit. In this case, we assumed the highest concentrations tested as surrogate IC_{50}. Such estimation may introduce additional variability, if different highest inhibitor concentrations are tested. It also cannot differentiate non vs. weak inhibition and makes the prediction not definitive. Furthermore, DDIs may be overestimated if the inhibitor concentration in gut is lower than I_2. This is consistent with our observations that 11 out of 15 FP prediction (model 1) are for class II or IV compounds (**Supplemental Table 5**). In the future, prediction of inhibitor concentration in gut using mechanistic modeling may improve the accuracy of DDI prediction. 2) Relevant inhibitor concentrations for IC_{50} measurement may be different from nominal concentrations. However, correction of nonspecific binding of inhibitor drugs in BDT assay did not improve the predictive performance. It is possible that unbound inhibitor concentrations measured in the incubation medium are not relevant concentrations for P-gp inhibition, as substrate binding sites of P-gp are localized intracellularly. Therefore, measuring intracellular unbound inhibitor concentrations and developing in vitro mechanistic modeling to determine true K_i values may improve the prediction of DDIs. 3) Mechanisms for P-gp inhibition have not been well characterized. It is not clear whether it is driven by cis- and /or trans-inhibition.

Understanding such mechanisms will help to improve in vitro to in vivo extrapolation. 4) It is worth noting that possible involvement of other mechanisms for the disposition of P-gp probe drugs, e.g., digoxin, DE and fexofenadine (Taub et al., 2011; Shimizu et al., 2005; Chu et al., 2018a; Medwid et al., 2019; Shen et al., 2019), may confound the DDI prediction, as they are not captured in such simplified P-gp inhibition models.

In conclusion, our studies confirmed that model 1 (I_2/IC_{50}) is sufficient to predict P-gp mediated intestinal DDIs. In evaluation of 7 P-gp IC_{50} measurement /calculation methods, IC_{50}(V), IC_{50}(NSF), and IC_{50}(ER) provided better predictive performance than others. Further analysis indicated that IC_{50}(V) obtained from MDR1 vesicles with refined threshold of $I_2/IC_{50} \geq 25.9$ provided comparable predictive power than BDT assays in LLC-MDR1 cells with IC_{50}(NSF) and IC_{50}(ER) at the threshold of $I_2/IC_{50} \geq 9.3$ and 13.7, respectively. These IC_{50} assays and cut-off values could also be used to predict P-gp mediated intestinal DDIs for DE and fexofenadine, despite additional data are needed for further validation. Through comprehensive calibration, our studies demonstrate that IC_{50} data generated from vesicular inhibition assay using NMQ as probe substrate are predictive for P-gp related intestinal DDIs with digoxin. We therefore recommend vesicular inhibition assay as our preferred method to study P-gp-mediated intestinal DDIs for digoxin (**Figure 5**), due to its simplicity, lower variability, higher assay throughput, and more direct estimation of kinetic parameters than BDT assays. This assay also offers the potential to predict DDIs for other P-gp probe substrates provided confirmatory validation is conducted.

Acknowledgement:

We are thankful to Lisa G Fox and J George Zhang from Corning Life Sciences for their expert assistance with experiments.

Authorship Contributions

Participated in research design: Chu, Yabut, Houle, Wang, Liaw

Conducted experiments: Yabut, Houle, Katwaru

Contributed new reagents or analytic tools: NA

Performed data analysis: Yabut, Houle, Wang, Chu, Collier

Wrote or contributed to the writing of the manuscript: Chu, Houle, Wang, Yabut, Liaw, Katwaru, Hittle

References

Agarwal S, Arya V, and Zhang L (2013) Review of P-gp inhibition data in recently approved new drug applications: utility of the proposed [I(1)]/IC(50) and [I(2)]/IC(50) criteria in the P-gp decision tree. *J Clin Pharmacol* **53:**228-233.

Balimane PV, Marino A, and Chong S (2008) P-gp inhibition potential in cell-based models: which "calculation" method is the most accurate? *AAPS J* **10:**577-586.

Benet LZ (2013) The role of BCS (biopharmaceutics classification system) and BDDCS (biopharmaceutics drug disposition classification system) in drug development. *J Pharm Sci* **102:**34-42.

Bentz J, O'Connor MP, Bednarczyk D, Coleman J, Lee C, Palm J, Pak YA, Perloff ES, Reyner E, Balimane P, Brannstrom M, Chu X, Funk C, Guo A, Hanna I, Heredi-Szabo K, Hillgren K, Li L, Hollnack-Pusch E, Jamei M, Lin X, Mason AK, Neuhoff S, Patel A, Podila L, Plise E, Rajaraman G, Salphati L, Sands E, Taub ME, Taur JS, Weitz D, Wortelboer HM, Xia CQ, Xiao G, Yabut J, Yamagata T, Zhang L, and Ellens H (2013) Variability in P-glycoprotein inhibitory potency (IC(5)(0)) using various in vitro experimental systems: implications for universal digoxin drug-drug interaction risk assessment decision criteria. *Drug Metab Dispos* **41:**1347-1366.

Brouwer KL, Keppler D, Hoffmaster KA, Bow DA, Cheng Y, Lai Y, Palm JE, Stieger B, Evers R, and International Transporter C (2013) In vitro methods to support transporter evaluation in drug discovery and development. *Clin Pharmacol Ther* **94:**95-112.

Chan G, Houle R, Lin M, Yabut J, Cox K, Wu J, and Chu X (2019) Role of transporters in the disposition of a novel beta-lactamase inhibitor: relebactam (MK-7655). *J Antimicrob Chemother* **74:**1894-1903.

Chu X, Galetin A, Zamek-Gliszczynski MJ, Zhang L, Tweedie DJ, and International Transporter C (2018a) Dabigatran Etexilate and Digoxin: Comparison as Clinical Probe Substrates for Evaluation of P-gp Inhibition. *Clin Pharmacol Ther* **104:**788-792.

Chu X, Liao M, Shen H, Yoshida K, Zur AA, Arya V, Galetin A, Giacomini KM, Hanna I, Kusuhara H, Lai Y, Rodrigues D, Sugiyama Y, Zamek-Gliszczynski MJ, Zhang L, and International Transporter C (2018b) Clinical Probes and Endogenous Biomarkers as Substrates for Transporter Drug-Drug Interaction Evaluation: Perspectives From the International Transporter Consortium. *Clin Pharmacol Ther* **104:**836-864.

Cook JA, Feng B, Fenner KS, Kempshall S, Liu R, Rotter C, Smith DA, Troutman MD, Ullah M, and Lee CA (2010) Refining the in vitro and in vivo critical parameters for P-glycoprotein, [I]/IC50 and [I2]/IC50, that allow for the exclusion of drug candidates from clinical digoxin interaction studies. *Mol Pharm* **7:**398-411.

Ellens H, Deng S, Coleman J, Bentz J, Taub ME, Ragueneau-Majlessi I, Chung SP, Heredi-Szabo K, Neuhoff S, Palm J, Balimane P, Zhang L, Jamei M, Hanna I, O'Connor M, Bednarczyk D, Forsgard M, Chu X, Funk C, Guo A, Hillgren KM, Li L, Pak AY, Perloff ES, Rajaraman G, Salphati L, Taur JS, Weitz D, Wortelboer HM, Xia CQ, Xiao G, Yamagata T, and Lee CA (2013) Application of receiver operating characteristic analysis to refine the prediction of potential digoxin drug interactions. *Drug Metab Dispos* **41:**1367-1374.

EMA C (2012) Guideline on the investigation of drug interactions European Medicines Agency. *https://wwwemaeuropaeu/documents/scientific-guideline/guideline-investigation-drug-interactions_enpdf*.

FDA C (2020) In Vitro Drug Interaction Studies — Cytochrome P450 Enzyme- and Transporter-Mediated Drug Interactions Guidance for Industry, in: *Center for Drug Evaluation and Research (CDER)*.

Fekete Z, Rajnai Z, Nagy T, Jakab KT, Kurunczi A, Gemes K, Heredi-Szabo K, Fulop F, Toth GK, Czerwinski M, Loewen G, and Krajcsi P (2015) Membrane Assays to Characterize Interaction of Drugs with ABCB1. *J Membr Biol* **248:**967-977.

Fenner KS, Troutman MD, Kempshall S, Cook JA, Ware JA, Smith DA, and Lee CA (2009) Drug-drug interactions mediated through P-glycoprotein: clinical relevance and in vitro-in vivo correlation using digoxin as a probe drug. *Clin Pharmacol Ther* **85:**173-181.

Hanley JA and McNeil BJ (1982) The meaning and use of the area under a receiver operating characteristic (ROC) curve. *Radiology* **143:**29-36.

Heredi-Szabo K, Palm JE, Andersson TB, Pal A, Mehn D, Fekete Z, Beery E, Jakab KT, Jani M, and Krajcsi P (2013) A P-gp vesicular transport inhibition assay - optimization and validation for drug-drug interaction testing. *Eur J Pharm Sci* **49:**773-781.

International Transporter C, Giacomini KM, Huang SM, Tweedie DJ, Benet LZ, Brouwer KL, Chu X, Dahlin A, Evers R, Fischer V, Hillgren KM, Hoffmaster KA, Ishikawa T, Keppler D, Kim RB, Lee CA, Niemi M, Polli JW, Sugiyama Y, Swaan PW, Ware JA, Wright SH, Yee SW, Zamek-Gliszczynski MJ, and Zhang L (2010) Membrane transporters in drug development. *Nat Rev Drug Discov* **9:**215-236.

Jani M and Krajcsi P (2014) In vitro methods in drug transporter interaction assessment. *Drug Discov Today Technol* **12:**e105-112.

Kishimoto W, Ishiguro N, Ludwig-Schwellinger E, Ebner T, Maeda K, and Sugiyama Y (2016) Usefulness of A Model-Based Approach for Estimating In Vitro P-Glycoprotein Inhibition Potency in a Transcellular Transport Assay. *J Pharm Sci* **105:**891-896.

Lappin G, Shishikura Y, Jochemsen R, Weaver RJ, Gesson C, Houston B, Oosterhuis B, Bjerrum OJ, Rowland M, and Garner C (2010) Pharmacokinetics of fexofenadine: evaluation of a microdose and assessment of absolute oral bioavailability. *Eur J Pharm Sci* **40:**125-131.

Lee CA, Cook JA, Reyner EL, and Smith DA (2010) P-glycoprotein related drug interactions: clinical importance and a consideration of disease states. *Expert Opin Drug Metab Toxicol* **6:**603-619.

Lee CA, Kalvass JC, Galetin A, and Zamek-Gliszczynski MJ (2014) ITC commentary on the prediction of digoxin clinical drug-drug interactions from in vitro transporter assays. *Clin Pharmacol Ther* **96:**298-301.

Li F, Howard KD, and Myers MJ (2017) Influence of P-glycoprotein on the disposition of fexofenadine and its enantiomers. *J Pharm Pharmacol* **69:**274-284.

Medwid S, Li MMJ, Knauer MJ, Lin K, Mansell SE, Schmerk CL, Zhu C, Griffin KE, Yousif MD, Dresser GK, Schwarz UI, Kim RB, and Tirona RG (2019) Fexofenadine and Rosuvastatin Pharmacokinetics in Mice with Targeted Disruption of Organic Anion Transporting Polypeptide 2B1. *Drug Metab Dispos* **47:**832-842.

Nervi P, Li-Blatter X, Aanismaa P, and Seelig A (2010) P-glycoprotein substrate transport assessed by comparing cellular and vesicular ATPase activity. *Biochim Biophys Acta* **1798:**515-525.

O'Connor M, Lee C, Ellens H, and Bentz J (2015) A novel application of t-statistics to objectively assess the quality of IC50 fits for P-glycoprotein and other transporters. *Pharmacol Res Perspect* **3:**e00078.

Papich MG and Martinez MN (2015) Applying Biopharmaceutical Classification System (BCS) Criteria to Predict Oral Absorption of Drugs in Dogs: Challenges and Pitfalls. *AAPS J* **17:**948-964.

Poirier A, Cascais AC, Bader U, Portmann R, Brun ME, Walter I, Hillebrecht A, Ullah M, and Funk C (2014) Calibration of in vitro multidrug resistance protein 1 substrate and inhibition assays as a basis to support the prediction of clinically relevant interactions in vivo. *Drug Metab Dispos* **42:**1411-1422.

Robin X, Turck N, Hainard A, Tiberti N, Lisacek F, Sanchez JC, and Muller M (2011) pROC: an open-source package for R and S+ to analyze and compare ROC curves. *BMC Bioinformatics* **12:**77.

Safar Z, Vasko B, Ritchie TK, Imre G, Mogyorosi K, Erdo F, Rajnai Z, Fekete Z, Szeremy P, Muka L, Zolnerciks JK, Heredi-Szabo K, Ragueneau-Majlessi I, and Krajcsi P (2016) Investigating ABCB1-Mediated Drug-Drug Interactions: Considerations for In vitro and In vivo Assay Design. *Curr Drug Metab* **17:**430-455.

Saito H, Osumi M, Hirano H, Shin W, Nakamura R, and Ishikawa T (2009) Technical pitfalls and improvements for high-speed screening and QSAR analysis to predict inhibitors of the human bile salt export pump (ABCB11/BSEP). *AAPS J* **11:**581-589.

Shen H, Yao M, Sinz M, Marathe P, Rodrigues AD, and Zhu M (2019) Renal Excretion of Dabigatran: The Potential Role of Multidrug and Toxin Extrusion (MATE) Proteins. *Mol Pharm* **16:**4065-4076.

Shimizu M, Fuse K, Okudaira K, Nishigaki R, Maeda K, Kusuhara H, and Sugiyama Y (2005) Contribution of OATP (organic anion-transporting polypeptide) family transporters to the hepatic uptake of fexofenadine in humans. *Drug Metab Dispos* **33:**1477-1481.

Sugimoto H, Matsumoto S, Tachibana M, Niwa S, Hirabayashi H, Amano N, and Moriwaki T (2011) Establishment of in vitro P-glycoprotein inhibition assay and its exclusion criteria to assess the risk of drug-drug interaction at the drug discovery stage. *J Pharm Sci* **100:**4013-4023.

Sziraki I, Erdo F, Beery E, Molnar PM, Fazakas C, Wilhelm I, Makai I, Kis E, Heredi-Szabo K, Abonyi T, Krizbai I, Toth GK, and Krajcsi P (2011) Quinidine as an ABCB1 probe for testing drug interactions at the blood-brain barrier: an in vitro in vivo correlation study. *J Biomol Screen* **16:**886-894.

Taub ME, Mease K, Sane RS, Watson CA, Chen L, Ellens H, Hirakawa B, Reyner EL, Jani M, and Lee CA (2011) Digoxin is not a substrate for organic anion-transporting polypeptide transporters OATP1A2, OATP1B1, OATP1B3, and OATP2B1 but is a substrate for a sodium-dependent transporter expressed in HEK293 cells. *Drug Metab Dispos* **39:**2093-2102.

Volpe DA, Hamed SS, and Zhang LK (2014) Use of different parameters and equations for calculation of IC(5)(0) values in efflux assays: potential sources of variability in IC(5)(0) determination. *AAPS J* **16:**172-180.

Wu CY and Benet LZ (2005) Predicting drug disposition via application of BCS: transport/absorption/ elimination interplay and development of a biopharmaceutics drug disposition classification system. *Pharm Res* **22:**11-23.

Zamek-Gliszczynski MJ, Lee CA, Poirier A, Bentz J, Chu X, Ellens H, Ishikawa T, Jamei M, Kalvass JC, Nagar S, Pang KS, Korzekwa K, Swaan PW, Taub ME, Zhao P, Galetin A, and International Transporter C (2013) ITC recommendations for transporter kinetic parameter estimation and translational modeling of transport-mediated PK and DDIs in humans. *Clin Pharmacol Ther* **94:**64-79.

Zhang L, Huang SM, Reynolds K, Madabushi R, and Zineh I (2018) Transporters in Drug Development: Scientific and Regulatory Considerations. *Clin Pharmacol Ther* **104:**793-796.

Zhou T, Arya V, and Zhang L (2019) Comparing Various In Vitro Prediction Methods to Assess the Potential of a Drug to Inhibit P-glycoprotein (P-gp) Transporter In Vivo. *J Clin Pharmacol* **59:**1049-1060.

Footnotes

This work received no external funding. All studies conducted in this manuscript were sponsored by Merck Sharp & Dohme Corp., a subsidiary of Merck & Co., Inc., Kenilworth, NJ, USA. No author has an actual or perceived conflict of interest with the contents of this article.

Figure Legends

Figure 1: Comparison of in vitro IC_{50} values of 48 compounds measured in bidirectional and vesicular transport inhibition assays. Panel A: IC_{50}(NSF) vs. IC_{50}(V); Panel B: $IC_{50}(NSF_c)$ vs. IC_{50}(V); C; IC_{50}(ER) vs. IC_{50}(V); D: $IC_{50}(ER_c)$ vs. IC_{50}(V); E: IC_{50}(UDF) vs. IC_{50}(V); F: $IC_{50}(UDF_c)$ vs. IC_{50}(V). The symbols in red, blue, green and black circle represent the compounds classified as BCS I, II, III, and IV compounds, respectively; while the symbols in opened diamond and square represent the compounds that are classified as either BCS I or II, and BCS II or IV, respectively. In Panels A-1, B-1, C-1, D-1, E-1, and F-1, the solid, dotted, and dashed lines represent the line of unity and 3-fold, and 5-fold differences, respectively. Panels A-2, B-2, C-2, D-2, E-2, and F-2 are correlation plots of respective data. Panels A-3, B-3, C-3, D-3, E-3, and F-3 are the correlation plots excluding the compounds in red circles (alogliptin and maraviroc in all figures (A-2 to F-2), and captopril in B-2, D-2, and F-2. In all correlation plots, solid and dotted lines represent the regression line and its 95% confidence interval, respectively. All IC_{50} values (mean ± SD) were shown in Table 3.

Figure 2: Comparison of IC_{50} values measured in bidirectional transport assays using different calculation methods. Panel A: IC_{50}(NSF) vs. IC_{50}(UDF); Panel B: IC_{50}(NSF) vs. IC_{50}(ER); Panel C: IC_{50}(UDF) vs. IC_{50}(ER). The symbols in red, blue, green and black circle represent the compounds classified as BCS I, II, III, and IV compounds, respectively; while the symbols in opened diamond and square represent the compounds that are classified as either BCS I or II, and BCS II or IV, respectively. In Panels A-1, B-1, C-1, the solid, dotted, and dashed lines represent the line of unity and 3-fold, and 5-fold differences, respectively. Panels A-2, B-2, and C-2 are correlation plots of respective data. Panels A-3, B-3, and C-3 are the

correlation plots excluding the compounds in red circles (captopril and maraviroc). In all correlation plots, solid and dotted lines represent the regression line and its 95% confidence interval, respectively. All IC_{50} values (mean ± SD) were shown in Table 3.

Figure 3: The AUC_{ROC} values obtained from ROC analysis of different static models using various IC_{50} measurement. Estimated AUC_{ROC} values and their 95% confidence intervals of various static models were shown in panel A (I_2/IC_{50}), panel B (I_1/IC_{50}, or I_2/IC_{50}), and panel C (I_{1u}/IC_{50}, or I_2/IC_{50}). In vitro IC_{50} values (IC_{50}(V), IC_{50}(ER), IC_{50}(NSF), IC_{50}(UDF), $IC_{50}(ER_c)$, $IC_{50}(NSF_c)$, $IC_{50}(UDF_c)$) were measured as described in the *Methods and Materials* section.

Figure 4: ROC curves of various static models using different IC_{50} estimation methods. ROC curves with different IC_{50} measurement and calculation methods were shown in panel A (model 1: I_2/IC_{50}), panel B (model 2: I_1/IC_{50}, or I_2/IC_{50}), and panel C (model 3: I_{1u}/IC_{50}, or I_2/IC_{50}).

Figure 5: Recommended workflow to evaluate intestinal P-gp inhibition. This workflow is only suitable when orally administered digoxin is used as a clinical P-gp probe. Additional calibration will be needed for dabigatran etexilate and other orally administered P-gp probe substrates. If digoxin is a co-medication, IC_{50}(NSF) or IC_{50}(ER) measured by BDT assay using digoxin as a probe substrate may be helpful in clinical study design and data interpretation.

[a]:The threshold of 25.9 is obtained based on calibration of vesicular P-gp inhibition assay using ROC analysis shown in this paper. The calibration of your own assay using similar approach is recommended.

Table 1. Clinical digoxin DDI data used in ROC analysis[a]

Inhibitor	BCS[b]	Dose (mg)	I_1 (μM)[c]	f_u[d]	I_2 (μM)	AUCR	$C_{max}R$	Clinical DDIs Y/N[e]
Alogliptin	III	25	0.44	0.80	295	1.00	0.94	N
Amiodarone	II	400	3.5	0.0002	2479	1.63	1.72	Y
		600	5.3	0.0002	3719	1.69	1.75	Y
		800	7.0	0.0002	4960	1.68	1.84	Y
Apixiban	III	20	0.94	0.13	174	0.90	0.92	N
Asunaprevir	II	200	0.49	0.03	1069	1.30	1.09	Y
Atorvastatin	II	10	0.013	0.02	72	1.00	1.10	N
		80	0.12	0.02	573	1.10	1.20	N
Azilsartan	IV	80	9.24	0.01	563	1.03	0.94	N
Bosentan	II	500	5.90	0.02	3626	0.87	0.94	N
Canagliflozin	IV	300	7.60	0.0125	2597	1.20	1.36	Y
Captopril	III	12.5	0.50	0.75	230	1.39	1.59	Y
Carvedilol (Male*)	II	6.5	0.13	0.05	62	1.24	1.00	N
Carvedilol (Female*)	II	6.5	0.13	0.05	62	1.56	1.38	Y
Clarithromycin	II	500	3.20	0.54	2674	1.46	1.75	Y
Clopidogrel	II	75	0.012	0.02	932	1.02	1.10	N
Daclatasvir	II	60	2.34	0.01	325	1.27	1.65	Y
Diltiazem	I	60	0.17	0.22	579	1.51	1.37	Y
		90	0.49	0.22	868	1.40	1.38	Y
		180	0.20	0.22	1737	1.50	1.37	Y
Dronedarone	II	400	0.19	0.02	2874	2.57	1.75	Y
Elagolix	III	200	1.23	0.2	1267	1.26	1.71	Y
Elbasvir	IV	50	0.14	0.001	227	1.11	1.47	Y
Eliglustat	I	100	0.18	0.38	989	1.37	1.68	Y
		150	0.27	0.38	1479	1.41	1.64	Y
Etravirine	IV	200	2.20	0.01	1838	1.18	1.19	N

Felodipine	II	2.5	0.007	0.004	26	1.49	1.35	Y
		5	0.013	0.004	52	1.11	1.38	Y
		10	0.034	0.004	104	1.16	1.39	Y
Fidaxomicin	IV	200	0.005	0.022	756	1.12	1.18	N
Flibanserin	II	100	1.07	0.02	1025	1.93	1.46	Y
Isradipine	II, IV	5	0.03	0.03	54	1.11	1.26	Y
Itraconazole	II	200	0.90	0.002	1134	1.68	1.34	Y
Ivacaftor	II, IV	150	13.90	0.02	1529	1.32	1.23	Y
Maraviroc	III	300	0.52	0.24	2336	1.02	1.09	N
Mibefradil	II	50	0.92	0.004	404	1.08	1.22	N
		100	1.85	0.004	807	1.07	1.25	Y
		150	2.77	0.004	1211	1.31	1.41	Y
Mirabegron	III	100	1.00	0.29	1009	1.27	1.29	Y
Nelfinavir	II	1250	7.04	0.015	8806	1.35	1.34	Y
Nicardipine	I	20	0.11	0.0125	167	1.07	NA	N
		30	0.17	0.0125	250	0.96	0.90	N
Nifedipine	I, II	5	0.12	0.04	58	1.21	1.01	N
		10	0.23	0.04	115	1.23	1.06	N
		20	0.46	0.04	231	1.18	1.08	N
Nitrendipine	II	10	0.01	0.02	111	1.09	1.22	N
		20	0.02	0.02	222	1.15	1.57	Y
Paroxetine	I	30	0.18	0.05	364	0.85	0.90	N
Quinidine	I	200	2.00	0.13	2466	NA	1.42	Y
		250	2.50	0.13	3082	NA	2.18	Y
		600	6.00	0.13	7398	2.65	NA	Y
Ranolazine	II	400	3.36	0.38	3742	1.39	2.30	Y
		750	6.68	0.38	7017	1.88	1.68	Y
		1000	8.40	0.38	9356	1.60	1.46	Y
Ritonavir	II	200	5.00	0.015	1110	1.22	1.26	Y
		400	10.00	0.015	2219	1.39	1.25	Y
Rolapitant	II	180	1.93	0.002	1439	1.30	1.71	Y
Rosuvastatin	II	40	0.02	0.12	332	1.04	1.01	N
Sertraline	I, II	200	0.41	0.015	2612	1.10	1.05	N
Telaprevir	II	750	5.24	0.325	4413	1.85	1.50	Y

Telmisartan	II	120	1.12	0.005	933	1.22	1.50	Y
Ticagrelor	IV	400	6.99	0.02	3062	1.28	1.75	Y
Troglitazone	II	400	3.00	0.01	3624	1.04	1.05	N
Valspodar	IV	200	1.49	0.025	659	3.05	2.44	Y
		400	1.56	0.025	1317	1.74	1.72	Y
Vandetanib	II	300	3.32	0.06	2524	1.23	1.29	Y
Velpatasvir	IV	100	0.70	0.005	453	1.27	1.87	Y
Vemurafenib	IV	960	125.30	0.01	7838	1.91	1.42	Y
Verapamil	I	80	0.37	0.1	704	1.50	1.44	Y
		120	0.55	0.1	1056	NA	1.61	Y

AUC, area under the concentration-time curve; BCS, biopharmaceutics classification; DDI, drug-drug interaction; IC_{50}, half maximal inhibitory concentration; I_1, maximum plasma concentration (C_{max}) of inhibitor drug following administration of clinical dose indicated; f_u, the fraction unbound in plasma; $I_{1,u}$, the unbound I_1; $I_{2,}$ the concentration of inhibitor drug in the gastrointestinal tract at indicated dose dissolved in 250 ml; NA, Not available.

[a] All clinical data in this table were used for ROC analysis as training set. [b]BCS classification of inhibitor drugs was obtained from the University of Washington DDI database (https://www.druginteractionsolutions.org), https://www.pharmapendium.com, and the literature (Wu and Benet, 2005; Benet, 2013; Papich and Martinez, 2015). [c] Clinical DDI data, I_1, and f_u values were collected from the University of Washington DDI database (https://www.druginteractionsolutions.org); I_1 reported from clinical DDI studies was used. Otherwise, I_1 at same or similar inhibitor dose regimen was collected. In the cases that I_1 at indicated inhibitor doses was not reported, extrapolated I_1 was used, assuming the linear PK. [d] Either AUCR and/or C_{max}R (pharmacokinetic (PK) ratios) greater or equal to 1.25 was considered a positive clinical DDI.

[d] If f_u measured < 0.01, f_u value of 0.01 was used for DDI prediction.

[e] Yes (Y) indicates AUCR or C_{max}R ≥1.25; No (N) indicates AUCR and C_{max}R <1.25

*Male and female cohort tested.

Table 2: Selected clinical P-gp DDI data using dabigatran etexilate and fexofenadine as probe drugs[a]

Inhibitor	BCS[b]	Inhibitor dose (mg)	I_1 (μM)[c]	I_2 (μM)	f_u [c]	Clinical probe[d]	Probe drug dose (mg)	AUCR[c]	C_{max}R[c]	Clinical DDIs (Y/N)[e]
Amiodarone	II	600	2.43	3720	0.002	DE	150	1.6	1.5	Y
Atorvastatin	II	80	0.12	573	0.026	DE	150	0.82	0.8	N
Clarithromycin	III	500	0.002	2674	0.5	DE	150	1.6	1.49	Y
Clarithromycin	III	500	3.2	2674	0.5	DE	0.375	4.22	4.58	Y
Clopidogrel	II	75	0.008	932	0.02	DE	150	0.92	0.95	N
Clopidogrel	II	300	0.055	3729	0.02	DE	150	1.36	1.68	Y
Clopidogrel	II	600	0.055	7458	0.02	DE	150	1.32	1.43	Y
Cobicistat	II	150	0.0015	773	0.053	DE	150	2.4	2.33	Y
Dronedarone	II	400	0.2	2874	0.02	DE	150	2.36	2.25	Y
Itraconazole	II	200	0.75	1134	0.002	DE	0.375	7.4	6.42	Y
Quinidine	I	200	0.56	2466	0.13	DE	150	1.53	1.56	Y
Rifampin	II	600	18.96	2916	0.4	DE	0.375	2.38	1.78	Y
Ritonavir	IV	100	2	555	0.31	DE	150	1.11	1.13	N
Ticagrelor	IV	90	0.899	689	0.016	DE	150	1.56	1.46	Y
Ticagrelor	IV	180	2.76	1378	0.016	DE	150	1.73	1.95	Y
Verapamil	II	120*	0.13	1056	0.225	DE	150	1.39, 2.08[f]	1.33, 2.29[f]	Y
Verapamil	II	240	0.26	2112	0.225	DE	150	1.71	1.91	Y
Alogliptin	III	100	2.058	1179	0.694	fexofenadine	80	1.26	1.07	Y
Itraconazole	II	200	0.483	1134	0.002	fexofenadine	180	2.29	2.69	Y
Paroxetine	I	20	0.02	4000	0.05	fexofenadine	60	1.38	1.33	Y
Quinidine	I	200	NA	2466	0.13	fexofenadine	25	2.14	2.39	Y
Sertraline	I	50	NA	653	0.01	fexofenadine	50	0.84	0.86	N
Verapamil	II	240	NA	2112	0.225	fexofenadine	120	1.46, 2.5[f]	1.3, 2.9[f]	Y

AUC, area under the concentration-time curve; BCS, biopharmaceutics classification; DDI, drug-drug interaction; DE, dabigatran etexilate; I_1, maximum plasma concentration (C_{max}) of inhibitor drug following administration of clinical dose indicated; f_u, the fraction unbound in plasma; I_2, the concentration of inhibitor drug in the gastrointestinal tract at indicated dose dissolved in 250 ml.

[a] Only clinical DDI data with inhibitor drug IC_{50} data shown in Table 1 were collected.

[b]BCS classification of inhibitor drugs was obtained from the University of Washington DDI database (https://www.druginteractionsolutions.org), https://www.pharmapendium.com, and the literature (Wu and Benet, 2005; Benet, 2013; Papich and Martinez, 2015).

[c] Clinical DDI data, I_1, and f_u values were collected from the University of Washington DDI database (https://www.druginteractionsolutions.org); I_1 reported from clinical DDI studies was used. Otherwise, I_1 at same or similar inhibitor dose regimen was collected. In the cases that I_1 at indicated inhibitor doses was not reported, extrapolated I_1 was used, assuming the linear PK. If f_u measured < 0.01, f_u value of 0.01 was used for DDI prediction.

DE, the prodrug of dabigatran was administered for dabigatran DDI studies.

[e]Either AUCR and/or $C_{max}R$ (pharmacokinetic (PK) ratios) greater or equal to 1.25 was considered a positive clinical DDI; Yes (Y) indicates AUCR or $C_{max}R \geq 1.25$; No (N) indicates AUCR and $C_{max}R < 1.25$.

[f] Multiple clinical DDI data reported.

Table 3: In vitro IC_{50} values of 48 compounds measured in P-gp bidirectional and vesicular transport inhibition assays[a]

Compounds	IC_{50}(NSF)	IC_{50}(NSF_c)	IC_{50}(ER)	IC_{50}(ER_c)	IC_{50}(UDF)	IC_{50}(UDF_c)	IC_{50}(V)
	(μM)						
Alogliptin	>100	>100	>100	>100	~ 207.0 ± 91.0	~ 207.0 ± 91.0	>1000
Amiodarone	21.9 ± 2.6	0.6 ± 0.1	8.3 ± 2.6	0.2± 0.1	10.6 ± 0.6	0.3 ± 0.02	14.8 ± 0.7
Apixaban	>150	>150	>150	>150	>150	>150	>150
Asunaprevir	21.2 ± 2.0	21.2 ± 2.2	17.3 ± 2.7	17.3± 2.7	23.7 ± 3.0	23.7 ± 3.0	3.2 ± 0.2
Atorvastatin	>100	>100	>100	>100	>100	>100	17.0 ± 1.7
Azilsartan	13.2 ± 2.7	2.6 ± 0.5	4.3 ± 0.3	0.8 ± 0.1	20.9 ± 2.5	4.2 ± 0.5	8.2 ± 0.7
Bosentan	>15	>15	>15	>15	>15	>15	>15
Canagliflozin	59.0 ± 10.2	2.9 ± 0.5	39.9 ± 9.0	2.0 ± 0.5	>100	> 5	>100
Captopril	>1000	>67	>1000	>67	>1000	>67	>1000
Carvedilol	19.1 ± 1.9	11.0 ± 0.9	1.6 ± 0.1	0.5 ± 0.04	16.4 ± 2.4	5.5 ± 0.8	4.1 ± 0.3
Clarithromycin	~145.7 ± 26.6	~145.7 ± 26.6	>100	>100	>100	>100	10.9 ± 2.4
Clopidogrel	>100	>36	132.8 ± 50.0	48.4 ± 18.3	>100	>36	157.4 ± 15.1
Daclatasvir	>100	>5	35.5 ± 6.9	1.8 ± 0.3	>100	>5	1.6 ± 0.1
Diltiazem	83.2 ± 14.8	83.2 ± 14.8	62.7 ± 11.6	62.7 ± 11.6	23.7 ± 7.2	23.7 ± 7.2	30.5 ± 2.8
Dronedarone	82.0 ± 29.0	3.1 ± 1.1	7.1 ± 1.6	0.3± 0.1	>300	>11	4.9 ± 0.5
Elagolix	~168.7 ± 51.4	~ 168.7 ± 51.4	100.5 ± 31.0	100.5 ± 31.0	>150	>150	24.8 ± 0.8
Elbasvir	>0.5	>0.5	>0.5	>0.5	>0.5	>0.5	0.3 ± 0.02
Eliglustat	>100	>100	43.7 ± 12.4	43.7 ± 12.4	>100	>100	65.4 ± 2.1
Etravirine	>10	>2.1	>10	>2.1	>10	>2.1	>30
Felodipine	11.4 ± 0.9	3.6 ± 0.2	3.8 ± 0.6	0.8 ± 0.01	>50	>10.5	81.8 ± 5.1
Fidaxomicin	>100	>100	>100	>100	>100	>100	0.4 ± 0.05
Flibanserin	51.4 ± 5.7	51.4 ± 5.7	8.8 ± 0.7	8.8 ± 0.7	~141.3 ± 76.3	~141.3 ± 76.3	>120
Isradipine	29.5 ± 1.7	29.5 ± 1.7	6.1 ± 0.3	6.1 ± 0.3	62.8 ± 16.6	62.8 ± 16.6	53.7 ± 2.4
Itraconazole	6.9 ± 0.7	1.2 ± 0.1	1.1 ± 0.3	0.2 ± 0.04	2.4 ± 0.2	0.4 ± 0.04	0.34 ± 0.13
Ivacaftor	1.8 ± 0.6	0.3 ± 0.1	2.1 ± 1.0	0.4 ± 0.2	0.6 ± 0.1	0.1 ± 0.04	> 1.0
Maraviroc	>1000	>1000	>1000	>1000	>1000	>1000	162.6 ± 8.5
Mibefradil	7.8 ± 1.8	1.9 ± 0.4	5.2 ± 0.4	1.3 ± 0.1	14.0 ± 5.2	3.4 ± 1.3	10.0 ± 1.6
Mirabegron	>300	>300	>300	>300	>300	>300	148.9 ± 21.8
Nelfinavir	14.8 ± 3.0	2.7 ± 0.5	1.3 ± 0.2	0.3 ± 0.03	7.3 ± 1.3	1.4 ± 0.2	20.3 ± 1.6
Nicardipine	5.3 ± 0.6	1.5 ± 0.2	0.8 ± 0.1	0.2 ± 0.02	4.1 ± 1.6	1.2 ± 0.5	6.4 ± 1.7
Nifedipine	83.8 ± 9.8	14.2 ± 2.6	18.3 ± 0.8	3.1 ± 0.1	58.9 ± 19.0	10.0 ± 3.2	115.6 ± 6.4
Nitrendipine	20.0 ± 4.1	20.0 ± 4.1	3.1 ± 0.1	3.1 ± 0.1	>100	> 100	76.0 ± 1.4
Paroxetine	>100	> 21.5	9.3 ± 0.9	2.0 ± 0.3	>100	> 21.5	122.2 ± 15.3
Quinidine	56.0 ± 9.5	56.0 ± 9.5	12.5 ± 2.2	12.5 ± 2.2	59.0 ± 12.8	59.0 ± 12.8	14.5 ± 1.9
Ranolazine	74.5 ± 9.6	74.5 ± 9.6	13.2 ± 0.7	13.2 ± 0.7	215.6 ± 62.0	215.6 ± 62.0	64.1 ± 4.1
Ritonavir	18.9 ± 1.0	18.9 ± 1.0	14.7 ± 0.6	14.7 ± 0.6	> 15	>15	0.3 ± 0.03
Rolapitant	22.2 ± 6.3	5.6 ± 0.4	6.7 ± 0.5	1.6 ± 0.1	34.9 ± 15.6	8.4 ± 3.7	>30
Rosuvastatin	>300	>300	>300	>300	>300	>300	>300
Sertraline	30.6 ± 12.8	7.7 ± 3.2	6.3 ± 1.0	1.6 ± 0.2	6.5 ± 0.8	1.7 ± 0.2	39.4 ± 2.8
Telaprevir	>200	>200	>200	>200	>200	>200	4.1 ± 0.5
Telmisartan	20.3 ± 2.8	20.3 ± 2.8	1.6 ± 0.2	1.6 ± 0.2	25.9 ± 8.9	25.9 ± 8.9	0.8 ± 0.1
Ticagrelor	~ 11.2 ± 3.1	~0.8 ± 0.3	3.7 ± 0.01	0.9 ± 0.002	~12.0 ± 6.1	~3.0 ± 1.5	>30
Troglitazone	17.1 ± 1.1	3.1 ± 0.2	9.3 ± 1.7	1.7 ± 0.3	13.5 ± 1.1	2.4 ± 0.2	19.4 ± 1.3
Valspodar	0.7 ± 0.1	0.1 ± 0.003	0.4 ± 0.1	0.04 ± 0.01	1.2 ± 0.3	0.05 ± 0.004	0.2 ± 0.01
Vandetanib	>10	>0.2	5.9 ± 0.9	0.1 ± 0.02	>10	>0.2	96.1 ± 9.1
Velpatasvir	>50	>2.5	>50	>2.5	>50	>2.5	4.9 ± 0.7
Vemurafenib	60.2 ± 25.0	8.8 ± 6.0	34.0 ± 15.0	5.3 ± 2.0	1.7 ± 0.9	0.3 ± 0.2	>30
Verapamil	39.6 ± 3.9	39.6 ± 3.9	3.2 ± 0.2	3.2 ± 0.2	52.9 ± 11.1	52.9 ± 11.1	2.8 ± 0.3

IC_{50}(NSF), IC_{50}(ER), and IC_{50}(UDF) represent IC_{50} values measured in bidirectional transport (BDT) assay estimated by calculating digoxin net transport using net secretory flux, efflux ratio, unidirectional flux, respectively. $IC_{50}(NSF_c)$, $IC_{50}(ER_c)$, and $IC_{50}(UDF_c)$, represent corrected IC_{50} values based on the recovery of inhibitor drugs measured in BDT assays. IC_{50}(V) represent IC_{50} values measured for inhibition of ATP-dependent NMQ vesicular transport. ~ (tilde operator) indicates approximation of IC_{50} extrapolated beyond maximum concentration tested. Data were reported as mean ± SD for triplicate measurement.

[a]: All data shown in this table were used as training set for ROC analysis.

Table 4. Summary of the cut-off values obtained from ROC analysis based on static P-gp DDI prediction models using various P-gp IC_{50} methods and the comparison of predictive performance with other cut-off criteria

Table 4-1: Model 1 (I_2/IC_{50})

Model 1 (I_2/IC_{50})	IC_{50}(V)[a]			IC_{50}(ER)[a]			IC_{50}(NSF)[a]		
	ROC Analysis	FDA 2000[b]	Ellens et al. 2013[c]	ROC Analysis	FDA 2000[b]	Ellens et al. 2013[c]	ROC Analysis	FDA 2000[b]	Ellens et al. 2013[c]
Cut-off values	25.9	10	45	13.7	10	45	9.3	10	45
TP % (sensitivity)	75 (36/48)	83.3 (40/48)	68.8 (33/48)	83.3 (40/48)	85.4 (41/48)	64.6 (31/48)	77.1 (37/48)	75 (36/48)	52.1 (25/48)
TN % (specificity)	59.1 (13/22)	45.5 (10/22)	72.7 (16/22)	50 (11/22)	45.5 (10/22)	63.6 (14/22)	63.6 (14/22)	63.6 (14/22)	72.7 (16/22)
FP %	40.9 (9/22)	54.5 (12/22)	27.3 (6/22)	50 (11/22)	54.5 (12/22)	36.4 (8/22)	36.4 (8/22)	36.4 (8/22)	27.3 (6/22)
FN %	25 (12/48)	16.7 (8/48)	31.3 (15/48)	16.7 (8/48)	14.6 (7/48)	35.4 (17/48)	22.9 (11/48)	25 (12/48)	47.9 (23/48)
Average Accuracy	0.67	0.644	0.707	0.667	0.654	0.641	0.704	0.693	0.624
Overall Accuracy	0.7	0.714	0.7	0.729	0.729	0.643	0.729	0.714	0.586

Table 4-2: Model 2 (I_1/IC_{50} or I_2/IC_{50})

Model 2 (I_1/IC_{50} or I_2/IC_{50})	IC_{50}(V)[a]			IC_{50}(ER)[a]			IC_{50}(NSF)[a]		
	ROC Analysis	Agarwal et al. 2013[d]	Ellens et al. 2013[c]	ROC Analysis	Agarwal et al. 2013[d]	Ellens et al. 2013[c]	ROC Analysis	Agarwal et al. 2013[d]	Ellens et al. 2013[c]
Cut-off values	(0.032, 40)	(0.1, 10)	(0.03, 45)	(0.081, 26.7)	(0.1, 10)	(0.03, 45)	(0.026, 10)	(0.1, 10)	(0.03, 45)
TP % (sensitivity)	75 (36/48)	83.3 (40/48)	75 (36/48)	75 (36/48)	85.4 (41/48)	70.8 (34/48)	75 (36/48)	75 (36/48)	54.2 (26/48)
TN % (specificity)	63.6 (14/22)	45.5 (10/22)	63.6 (14/22)	50 (11/22)	45.5 (10/22)	59.1 (13/22)	63.6 (14/22)	63.6 (14/22)	68.2 (15/22)
FP %	36.4 (8/22)	54.5 (12/22)	36.4 (8/22)	50 (11/22)	54.5 (12/22)	40.9 (9/22)	36.4 (8/22)	36.4 (8/22)	31.8 (7/22)
FN %	25 (12/48)	16.7 (8/48)	25 (12/48)	25 (12/48)	14.6 (7/48)	29.2 (14/48)	25 (12/48)	25 (12/48)	45.8 (22/48)
Average Accuracy	0.693	0.644	0.693	0.625	0.654	0.65	0.693	0.693	0.612
Overall Accuracy	0.714	0.714	0.714	0.671	0.729	0.671	0.714	0.714	0.586

Table 4-3: Model 3 (I_{1u}/IC_{50} or I_2/IC_{50})

Model 3 (I_{1u}/IC_{50} or I_2/IC_{50})	IC_{50}(V)[a]		IC_{50}(ER)[a]		IC_{50}(NSF)[a]	
	ROC Analysis	EMA 2012[e]	ROC Analysis	EMA 2012[e]	ROC Analysis	EMA 2012[e]
Cut-off values	(0.00141, 3334)	(0.02,10)	(0.00177,62)	(0.02,10)	(0.00052,94)	(0.02,10)
TP % (sensitivity)	75 (36/48)	83.3 (40/48)	75 (36/48)	85.4 (41/48)	75 (36/48)	75 (36/48)
TN % (specificity)	81.8 (18/22)	45.5 (10/22)	54.5 (12/22)	45.5 (10/22)	68.2 (15/22)	63.6 (14/22)
FP %	18.2 (4/22)	54.5 (12/22)	45.5 (10/22)	54.5 (12/22)	31.8 (7/2+2)	36.4 (8/22)
FN %	25 (12/48)	16.7 (8/48)	25 (12/48)	14.6 (7/48)	25 (12/48)	25 (12/48)
Average Accuracy	0.784	0.644	0.648	0.654	0.716	0.693
Overall Accuracy	0.771	0.714	0.686	0.729	0.729	0.714

EMA, European Medicines Agency; FDA, US Food and Drug Administration; TN, true negative; TP, true positive; FN, false

negative; FP, false positive; IC_{50}, half-maximal inhibitory concentration; $I_{1,}$ the mean steady-state total (free and bound) maximum plasma concentration (C_{max}) of inhibitor following administration of the highest proposed clinical dose; $I_{1,u,}$ unbound I_1; $I_{2,}$ the concentration of inhibitor in the gastrointestinal tract based on highest approved dose dissolved in 250 ml.

[a]: IC_{50}(V), IC_{50}(ER), and IC_{50}(NSF) were used in the static models for DDI prediction, respectively. IC_{50}(V), IC_{50}(ER), and IC_{50}(NSF) were determined as described in *Materials and Methods* section and shown in Table 3.

[b]: The cut-off value was obtained from FDA final DDI guidance (FDA 2020).

[c]: The cut-off value was obtained from Ellens et al (2013) based on ROC analysis of P-gp IC_{50} data for 15 compounds generated by 23 laboratories using four in vitro systems: Caco-2 cells, LLCPK1-MDR1, MDCKII-MDR1, and MDR1 vesicles. P-gp probe substrates were digoxin for polarized cell-lines and NMQ or vinblastine for MDR1 vesicles.

[d]: The cut-off value was obtained from FDA DDI draft guidance (Agarwal et al., 2013).

[e]: The cut-off value was obtained from EMA DDI guidance (EMA 2012).

Table 5. Summary of predictive performance of selected P-gp clinical inhibition studies using dabigatran etexilate and fexofenadine as in vivo probes and the cut-off values derived from ROC analysis based on digoxin clinical DDI data and the comparison with other cut-off criteria

Model 1 (I_2/IC_{50})	IC_{50}(V)[a]			IC_{50}(ER)[a]			IC_{50}(NSF)[a]		
	ROC Analysis[b]	FDA 2000[c]	Ellens et al. 2013[d]	ROC Analysis[b]	FDA 2000[c]	Ellens et al. 2013[d]	ROC Analysis[b]	FDA 2000[c]	Ellens et al. 2013[d]
Cut-off values	25.9	10.0	45.0	13.7	10.0	45.0	9.3	10.0	45.0
TP % (sensitivity)	84.2 (16/19)	94.7 (18/19)	78.9 (15/19)	89.5 (17/19)	94.7 (18/19)	73.7 (14/19)	100 (19/19)	100 (19/19)	47.4 (9/19)
TN % (specificity)	50 (2/4)	25 (1/4)	75 (3/4)	50 (2/4)	50 (2/4)	75 (3/4)	25 (1/4)	50 (2/4)	100 (4/4)
FP %	50 (2/4)	75 (3/4)	25 (1/4)	50 (2/4)	50 (2/4)	25 (1/4)	75 (3/4)	50 (2/4)	0 (0/4)
FN %	15.8 (3/19)	5.3 (1/19)	21.1 (4/19)	10.5 (2/19)	5.3(1/19)	26.3 (5/19)	0 (0/19)	0 (0/19)	52.6 (10/19)
Average Accuracy	0.783	0.826	0.783	0.826	0.870	0.739	0.870	0.913	0.565
Overall Accuracy	0.671	0.599	0.770	0.697	0.724	0.743	0.625	0.750	0.737

FDA, US Food and Drug Administration; TN, true negative; TP, true positive; FN, false negative; FP, false positive; IC_{50}, half-maximal inhibitory concentration; I_2, the concentration of inhibitor in the gastrointestinal tract based on highest approved dose dissolved in 250 ml.

[a]: IC_{50}(NSF), IC_{50}(ER), and IC_{50}(V) were used in the static models for DDI prediction, respectively. IC_{50}(NSF), IC_{50}(ER), and IC_{50}(V) of respective inhibitor drugs were shown in Table 3.

[b]: Cut-off values derived from ROC analysis based on digoxin clinical DDI data in Table 1 and IC_{50} data in Table 3.

[c]: The cut-off value was obtained from FDA final DDI guidance (FDA 2020).

[d]: The cut-off value was obtained from Ellens et al (2013) based on ROC analysis of P-gp IC_{50} data for 15 compounds generated by 23 laboratories using four in vitro systems: Caco-2 cells, LLCPK1-MDR1, MDCKII-MDR1, and MDR1 vesicles. P-gp probe substrates were digoxin for polarized cell-lines and NMQ or vinblastine for MDR1 vesicles.

Supplementary Tables for

Selection of an optimal in vitro model to assess P-gp inhibition: comparison of vesicular and bi-directional transcellular transport inhibition assays

Jocelyn Yabut, Robert Houle, Shubing Wang, Andy Liaw, Ravi Katwaru, Hannah Collier, Lucinda Hittle, and Xiaoyan Chu

Supplemental Table 1. LC/MS/MS Parameters

	MS Parameters[a]			
Compound	**Q1**	**Q3**	**CE**	**DP**
Alogliptin	340.2	116	38	105
Amiodarone	646.3	71.9	70	50
Apixaban	460.1	281.93	45	125
Asunaprevir	748.4	648.2	40	90
Atorvastatin	559.4	440.3	30	146
Azilsartan	569.3	233.1	55	90
Bosentan	552.3	202.1	60	115
Canagliflozin	445.3	267.2	25	65
Captopril	218.3	116	20	38
Carvedilol	407.5	224.1	30	100
Clarithromycin	748.6	158.2	50	53
Clopidogrel	322.1	155.1	45	45
Daclatasvir	739.5	565.2	55	200
Diltiazem	415	108.9	70	106
Dronedarone	646.3	58.1	70	50
Elagolix	632.3	529.3	65	160
Elbasvir	882.5	708.4	67	200
Eliglustat	405.3	149.1	15	135
Etravirine	435	303.8	52	200
Felodipine	384.2	338	24	65
Fidaxomicin	1055.7	231.1	-47	-100
Flibanserin	391.3	133.1	25	110
Isradipine	370.2	328.2	28	105
Itraconazole	705.5	392.2	50	56
Ivacaftor	393.3	115.9	100	64
Maraviroc	514.2	116.96	90	40
Mibefradil	496.4	202.1	65	89
Mirabegron	397.2	102.9	80	85
Nelfinavir	568.5	134.9	50	52
Nicardipine	480.3	315.1	50	90
Nifedipine	338.1	172.8	73	200

Nitrendipine	359.3	121.9	-20	-80
Paroxetine	329.8	70.1	110	45
Quinidine	325.2	307.1	70	30
Ranolazine	428.2	98.1	40	110
Ritonavir	721.5	296.2	25	85
Rolapitant	501.3	198.1	75	35
Rosuvastatin	482.4	270.19	52	120
Sertraline	306.1	158.9	90	35
Telaprevir	680.6	322.4	40	47
Telmisartan	515.6	305.2	30	131
Ticagrelor	523.3	153.1	55	120
Troglitazone	442.2	165.1	52	180
Valspodar	1215.7	425.3	59	90
Vandetanib	475.3	112.1	24	130
Velpatasvir	881.7	849.7	-30	-91
Vemurafenib	490.2	383.1	35	140
Verapamil	883.6	851.5	55	200

[a]The optimal MS parameters used for LC-MS/MS analysis

Supplemental Table 2: Clinical DDI data of test set compounds with orally administered digoxin

Inhibitor drugs	BCS [a]	Dose (mg)	I_1 [b] (μM)	I_2 (μM)	f_u [b, d]	AUC Ratio[b]	C_{max} Ratio[b]	Clinical DDI Y/N[c]
Cobicistat	II	150	2.34	773	0.03	1.09	1.47	Y
Dapagliflozin	III	10	0.35	98	0.09	1.00	0.99	N
Lapatinib	IV	1500	5.04	10326	0.01	1.63	2.09	Y
Linagliptin	III	5	0.01	42	0.11	1.01	0.94	N
Lurasidone	II	120	0.24	974	0.01	1.11	1.09	N
Neratanib	IV	240	0.13	1723	0.01	1.32	1.54	Y
Rifampin	II	600	8.00	2916	0.25	1.46	1.49	Y
Simeprevir	IV	150	10.88	800	0.001	1.39	1.31	Y
Tetrabenazine	III, IV	25	0.001	315	0.17	1.02	1.12	N
Tolvaptan	II, IV	60	0.96	535	0.02	1.18	1.27	Y
Valbenazine	I, III	80	2.19	765	0.01	1.33	1.87	Y
Vorapaxar	II	40	0.92	325	0.002	1.05	1.54	Y

AUC, area under the curve; BCS, biopharmaceutics classification; DDI, drug-drug interaction; I_1, maximum plasma concentration (C_{max}) of inhibitor drug following administration of clinical dose indicated; I_2, the concentration of inhibitor drug in the gastrointestinal tract at indicated dose dissolved in 250 ml; f_u, the fraction unbound in plasma.
[a] BCS classification of inhibitor drugs was obtained from the University of Washington DDI database, https://www.druginteractionsolutions.org, https://www.pharmapendium.com and the literature (Wu and Benet, 2005; Benet, 2013; Papich and Martinez, 2015). [b] Clinical DDI data, I_1, and f_u values were collected from the University of Washington DDI database (https://www.druginteractionsolutions.org). I_1 reported from clinical DDI studies was used. Otherwise, I_1 at same or similar inhibitor dose regimen was collected. In the cases that I_1 at indicated inhibitor doses was not reported, extrapolated I_1 was used, assuming the linear PK. [c] Either AUC and/or C_{max} ratios (pharmacokinetic (PK) ratios) greater or equal to 1.25 was considered a positive clinical DDI. Yes (Y) indicates AUC or C_{max} Ratio ≥1.25; No (N) indicates AUC and C_{max} Ratio <1.25; [d] If f_u measured < 0.01, f_u value of 0.01 was used for DDI prediction.

Supplemental Table 3: In vitro P-gp IC_{50} variability for verapamil and quinidine measured in MDR1 vesicles and bidirectional transport (BDT) inhibition assays

Inhibitor drugs	Vesicular assay (MDR1 vesicles)					BDT assay (LLC-MDR1 cells)				
		IC_{50}(V) (µM)[a]	IC_{50}(µM) (Mean ± SD)	CV %	IC_{50}(µM) reported in the literature[b]		IC_{50}(NSF) (µM)[c]	IC_{50}(µM) (Mean ± SD)	CV %	IC_{50}(µM) reported in the literature[d]
Verapamil	Study 1	3.2 ± 0.42	3.9 ± 0.7	18.2	1.2-3.3	Study 1	36.5 ± 5.0	45.5 ± 12.2	26.9	8.5-57
	Study 2	3.8 ± 0.17				Study 2	40.5 ± 3.5			
	Study 3	4.6 ± 0.48				Study 3	59.4 ± 7.7			
Quinidine	Study 1	6.4 ± 0.42	7.1 ± 0.6	8.9	1.0-9.8	Study 1	53.6 ± 3.5	48.0 ± 6.5	13.5	1.0-56
	Study 2	7.5 ± 0.62				Study 2	40.9 ± 4.4			
	Study 3	7.5 ± 0.15				Study 3	49.5 ± 3.0			

[a]: IC_{50}(V) was measured in MDR1 vesicles using [^{3}H]NMQ as an in vitro probe as described in the *Materials and Methods*.
[b]: IC_{50} values measured in MDR1 vesicles were collected from Ellens et al. (2013) using NMQ and /or vinblastine as in vitro probes and the University of Washington DDI database (https://www.druginteractionsolutions.org) using NMQ as an in vitro probe.
c: IC_{50}(NSF) was measured in LLCPK1-MDR1 cells using [^{3}H] digoxin as an in vitro probe as described in the *Materials and Methods*.
[b]: IC_{50} values measured in LLC-MDR1 cells were collected from Ellens et al. (2013) and the University of Washington DDI database (https://www.druginteractionsolutions.org) using digoxin as an in vitro probe.

Supplemental Table 4. Comparison of In vitro IC_{50}(V) values with those reported in the literature using MDR1 vesicular inhibition assays

		Reported in Literature		
Inhibitor	IC_{50}(V) (μM)[a]	IC_{50}(V) (μM)[b]	In Vitro Probe	Source
Alogliptin	>1000	NR		
Amiodarone	14.8 ± 0.7	1.6-9.3	NMQ/VB	UWDIDB; Ellens et al., 2013; Fekete et al., 2015
Apixaban	>150	NR		
Asunaprevir	3.2 ± 0.2	NR		
Atorvastatin	17.0 ± 1.7	23.7	NMQ	UWDIDB; Safar et al., 2018
Azilsartan	8.2 ± 0.7	NR		
Bosentan	>15	NR		
Canagliflozin	>100	NR		
Captopril	>1000	>1000	NMQ/VB	Bentz et al., 2013
Carvedilol	4.1 ± 0.3	0.6-3.4	NMQ/VB	UWDIDB; Ellens et al., 2013
Clarithromycin	10.9 ± 2.4	8.9	NMQ	UWDIDB; Vermeer et al., 2016
Clopidogrel	157.4 ± 15.1	NR		
Daclatasvir	1.6 ± 0.1	NR		
Diltiazem	30.5 ± 2.8	12.0-53.8	NMQ/VB	UWDIDB; Ellens et al., 2013
Dronedarone	4.9 ± 0.5	NR		
Elagolix	24.8 ± 0.8	NR		
Elbasvir	0.3 ± 0.02	0.32	NMQ	UWDIDB; NDA 208261
Eliglustat	65.4 ± 2.1	NR		
Etravirine	>30	NR		
Felodipine	81.8 ± 5.1	5.4-24.1	NMQ/VB	UWDIDB; Ellens et al., 2013
Fidaxomicin	0.4 ± 0.05	NR		
Flibanserin	>120	NR		
Isradipine	53.7 ± 2.4	4.6-16.4	NMQ/VB	UWDIDB; Ellens et al., 2013

Itraconazole	0.34 ± 0.13	0.048-2	NMQ	UWDIDB; Vermeer et al., 2016; Lempers et al., 2016
Ivacaftor	> 1.0	NR		
Maraviroc	162.6 ± 8.5	NR		
Mibefradil	10.0 ± 1.6	3.2-9.5	NMQ/VB	UWDIDB; Ellens et al., 2013
Mirabegron	148.9 ± 21.8	NR		
Nelfinavir	20.3 ± 1.6	2.72	NMQ	UWDIDB; Heredi-Szabo et al., 2013
Nicardipine	6.4 ± 1.7	0.7-3.7	NMQ/VB	UWDIDB; Ellens et al., 2013
Nifedipine	115.6 ± 6.4	10.7-54.7	NMQ/VB	UWDIDB; Ellens et al., 2013
Nitrendipine	76.0 ± 1.4	6.5-28.1	NMQ/VB	UWDIDB; Ellens et al., 2013
Paroxetine	122.2 ± 15.3	NR		
Quinidine	14.5 ± 1.9	1.0-9.8	NMQ/VB	UWDIDB; Ellens et al., 2013; Horio et al., 1988; Fekete et al., 2015
Ranolazine	64.1 ± 4.1	2.9-85.5	NMQ/VB	UWDIDB; Ellens et al., 2013
Ritonavir	0.3 ± 0.03	0.24-0.33	NMQ	UWDIDB; Heredi-Szabo et al., 2013; Vermeer et al., 2016
Rolapitant	>30	NR		
Rosuvastatin	>300	>300	NMQ	UWDIDB; Safar et al., 2018
Sertraline	39.4 ± 2.8	6.0-50.9	NMQ/VB	UWDIDB; Ellens et al., 2013
Telaprevir	4.1 ± 0.5	7	VC	UWDIDB; Fujita et al., 2013
Telmisartan	0.8 ± 0.1	0.1-3.6	NMQ/VB	UWDIDB; Ellens et al., 2013
Ticagrelor	>30	NR		
Troglitazone	19.4 ± 1.3	5.4-12.4	NMQ/VB	UWDIDB; Ellens et al., 2013
Valspodar	0.2 ± 0.01	0.031	NMQ	UWDIDB; Heredi-Szabo et al., 2013
Vandetanib	96.1 ± 9.1	NR		
Velpatasvir	4.9 ± 0.7	NR		
Vemurafenib	>30	NR		
Verapamil	2.8 ± 0.3	1.2-59	NMQ/VB/VC	UWDIDB; Ellens et al., 2013; Horio et al., 1988; Schaefer et al., 2006

UWDIDB: University of Washington DDI database (https://www.druginteractionsolutions.org); NR, not reported; NMQ, N-methyl-quinidine; VB, vinblastine; VC, vincristine.

[a]: IC_{50}(V) was measured in MDR1 vesicles using [^{3}H] NMQ as an in vitro probe as described in the *Materials and Methods*.

[b]: IC_{50}(V) values measured in MDR1 vesicles were collected from the University of Washington DDI database (https://www.druginteractionsolutions.org) and selected references using in vitro probe substrate(s) as indicated.

Supplemental Table 5. List of false negative and false positive prediction of digoxin clinical DDIs for training set compounds

		Model 1 (I_2/IC_{50})									Model 2 (I_1/IC_{50} or I_2/IC_{50})									Model 3 (I_{1u}/IC_{50} or I_2/IC_{50})					
		$IC_{50}(V)$[a]			$IC_{50}(ER)$[a]			$IC_{50}(NSF)$[a]			$IC_{50}(V)$[a]			$IC_{50}(ER)$[a]			$IC_{50}(NSF)$[a]			$IC_{50}(V)$[a]		$IC_{50}(ER)$[a]		$IC_{50}(NSF)$[a]	
		ROC Analysis	FDA 2000[b]	Ellens et al.[c]	ROC Analysis	FDA 2000[b]	Ellens et al.[c]	ROC Analysis	FDA 2000[b]	Ellens et al.[c]	ROC Analysis	Agarwal et al.[d]	Ellens et al.[c]	ROC Analysis	Agarwal et al[d]	Ellens et al.[c]	ROC Analysis	Agarwal et al.[d]	Ellens et al.[c]	ROC Analysis	EMA 2012[e]	ROC Analysis	EMA 2012[e]	ROC Analysis	EMA 2012[e]
	cut-off values	25.9	10	45	13.7	10	45	9.3	10	45	(0.032,40)	(0.1,10)	(0.03,45)	(0.081,26.7)	(0.1,10)	(0.03,45)	(0.026,10)	(0.1,10)	(0.03,45)	(0.00141,3334)	(0.02,10)	(0.00177,62)	(0.02,10)	(0.00052,94)	(0.02,10)
False negative	Asunaprevir, BCS II																					√			
	Canagliflozin, BCS IV			√																√					
	Captopril, BCS III	√	√	√	√	√	√	√	√	√	√	√	√	√	√	√	√	√	√	√	√	√	√	√	√
	Carvedilol, female, BCS II	√		√			√	√	√	√							√	√	√					√	√
	Clarithromycin, BCS II						√			√									√						
	Daclatasvir, BCS II				√	√	√	√	√	√				√	√		√	√	√			√	√	√	√
	Diltiazem 180mg, BCS I						√			√						√			√			√			
	Diltiazem 60mg, BCS I	√		√	√	√	√	√	√	√	√		√	√	√	√	√	√	√			√	√		√
	Diltiazem 90mg, BCS I			√			√			√	√		√	√		√			√			√			
	Dronedarone, BCS II									√									√	√				√	
	Elagolix, BCS III				√		√	√	√	√				√		√	√	√	√						√
	Eliglustat Tartrate 100mg, BCS I	√		√			√		√	√	√		√	√		√	√	√	√	√		√			√
	Eliglustat Tartrate, 150mg, BCS I	√		√			√			√	√		√			√			√						
	Felodipine 5mg, BCS II	√	√	√			√	√	√	√	√	√	√	√		√	√	√	√	√	√	√		√	√
	Felodipine 10mg, BCS II	√	√	√			√	√	√	√	√	√	√			√	√	√	√	√	√	√		√	√
	Felodipine 2.5mg, BCS II	√	√	√	√	√	√	√	√	√	√	√	√	√	√	√	√	√	√	√	√	√	√	√	√
	Flibanserin, BCS II	√	√	√						√	√	√	√						√	√	√			√	
	Isradipine, BCS II, IV	√	√	√	√	√	√	√	√	√	√	√	√	√	√	√	√	√	√	√	√	√	√	√	√
	Mirabegron, BCS III	√	√	√	√	√	√	√	√	√	√	√	√	√	√	√	√	√	√		√	√	√		√
	Nitrendipine 20mg, BCS II	√	√	√						√	√	√	√						√	√	√			√	
	Quinidine 200mg, BCS I									√															
	Rolapitant, BCS II																			√					
	Telaprevir, BCS II						√			√				√		√			√						
	Vandetanib, BCS II			√																					
	Velpatasvir, BCS IV				√	√	√	√	√	√				√	√	√	√	√	√			√	√	√	√

	Verapamil 120mg, BCS I									√									√					
	Verapamil 80mg, BCS I									√									√					
False positive	Alogliptin, BCS III																					√		√	
	Apixiban, BCS III																							√	
	Atorvastatin 80mg, BCS II	√	√									√									√				
	Azilsartan kamedoxomil, BCS IV	√	√	√	√	√	√	√	√		√	√	√	√	√	√	√	√	√	√	√	√	√	√	√
	Bosentan, BCS II	√	√	√	√	√	√	√	√	√	√	√	√	√	√	√	√	√	√	√	√	√	√	√	√
	Carvedilol, male, BCS II		√		√	√					√	√	√	√	√					√	√	√	√		
	Etravirine, BCS IV	√	√	√	√	√	√	√	√	√	√	√	√	√	√	√	√	√	√		√	√	√	√	√
	Fidaxomicin, BCS IV	√	√	√							√	√	√								√				
	Maraviroc, BCS III		√									√									√				
	Mibefradil 50mg, BCS II	√	√		√	√	√	√	√	√	√	√		√	√	√	√	√	√		√	√	√	√	√
	Nicardipine 20mg, BCS I	√	√		√	√	√	√	√			√		√	√	√	√	√			√	√	√		√
	Nicardipine 30mg, BCS I	√	√		√	√	√	√	√	√		√		√	√	√	√	√	√		√	√	√		√
	Nifedipine 20mg, BCS I, II					√									√								√		
	Nitrendipine 10mg, BCS II				√	√								√	√								√		
	Paroxetine, BCS I				√	√								√	√								√		
	Sertraline, BCS I, II	√	√	√	√	√	√	√	√	√	√	√	√	√	√	√	√	√	√		√	√	√	√	√
	Troglitazone, BCS II	√	√	√	√	√	√	√	√	√	√	√	√	√	√	√	√	√	√	√	√	√	√	√	√

[a]: IC_{50}(V), IC_{50}(ER), and IC_{50}(NSF) were used in the static models for DDI prediction, respectively. IC_{50}(V), IC_{50}(ER), and IC_{50}(NSF) was determined as described in the *Materials and Methods*.
[b]: The cut-off value was obtained from FDA final DDI guidance (FDA, 2020)
[c]: The cut-off value was obtained from Ellens et al (2013).
[d]: The cut-off value was obtained from (Agarwal et al., 2013)
[e]: The cut-off value was obtained from EMA DDI guidance (EMA, 2012)

Supplemental Table 6: In vitro IC_{50} values of test set compounds measured in bidirectional and vesicular transport inhibition assays

Inhibitor drugs	IC_{50}(NSF)[a] (μM)	IC_{50}(ER)[a] (μM)	IC_{50}(V)[b] (μM)
Cobicistat	47.0 ± 11.1	>100	0.3±0.01
Dapagliflozin	>100	>100	>100
Lapatinib	41.5 ± 8.5	17.2±4.7	>3
Linagliptin	>100	>100	62.7±7.6
Lurasidone	3.7 ± 0.5	1.9±0.6	3.2±0.6
Neratanib	31.1 ± 14.0	1.4±0.2	3.1±0.2
Rifampin	>30	>30	27.8±2.3
Simeprevir	5.6±0.4	5.1±0.4	5.9±0.5
Tetrabenazine	>100	37.0±15.2	>100
Tolvaptan	38.9 ± 2.2	17±5.2	18.8±1.3
Valbenazine	14.3± 1.7	3.7±0.3	8.4±0.6
Vorapaxar	5.9 ±0.8	2.7±0.3	7.6±0.6

[a]IC_{50}(NSF) and IC_{50}(ER) represent IC_{50} values estimated by calculating digoxin net transport using net secretory flux and efflux ratio, respectively. [b]IC_{50}(V) represent IC_{50} values measured for inhibition of ATP-dependent NMQ vesicular uptake. Data were reported as mean ± SD (n=3).

Supplemental Table 7. Summary of predictive performance of test set compounds using the cut-off values derived from ROC analysis and the comparison with other recommended cut-off criteria

Sup. Table 7-1: Model 1 (I_2/IC_{50})

Model 1 (I_2/IC_{50})	IC_{50}(V)[a]			IC_{50}(ER)[a]			IC_{50}(NSF)[a]		
	ROC Analysis	FDA 2020[d]	Ellens et al. 2013[c]	ROC Analysis	FDA 2020[d]	Ellens et al. 2013[c]	ROC Analysis	FDA 2000[b]	Ellens et al. 2013[c]
Cut-off values	25.9	10	45	13.7	10	45	9.3	10	45
TP % (sensitivity)	100 (8/8)	100 (8/8)	75 (6/8)	87.5 (7/8)	87.5 (7/8)	75 (6/8)	100 (8/8)	100 (8/8)	75 (6/8)
TN % (specificity)	75 (3/4)	75 (3/4)	75 (3/4)	75 (3/4)	75 (3/4)	75 (3/4)	75 (3/4)	75 (3/4)	75 (3/4)
FP %	25 (1/4)	25 (1/4)	25 (1/4)	25 (1/4)	25 (1/4)	25 (1/4)	25 (1/4)	25 (1/4)	25 (1/4)
FN %	0 (0/8)	0 (0/8)	25 (2/8)	12.5 (1/8)	12.5 (1/8)	25 (2/8)	0 (0/8)	0 (0/8)	25 (2/8)
Average Accuracy	0.917	0.917	0.75	0.833	0.833	0.75	0.917	0.917	0.75
Overall Accuracy	0.875	0.875	0.75	0.812	0.812	0.75	0.875	0.875	0.75

Sup. Table 7-2: Model 2 (I_1/IC_{50} or I_2/IC_{50})

Model 2 (I_1/IC_{50} or I_2/IC_{50})	IC_{50}(V)[a]			IC_{50}(ER)[a]			IC_{50}(NSF)[a]		
	ROC Analysis	Agarwal et al 2013[b]	Ellens et al. 2013[c]	ROC Analysis	Agarwal et al 2013[b]	Ellens et al. 2013[c]	ROC Analysis	Agarwal et al 2013[b]	Ellens et al. 2013[c]
Cut-off values	(0.032,40)	(0.1,10)	(0.03,45)	(0.081,26.7)	(0.1,10)	(0.03,45)	(0.026,10)	(0.1,10)	(0.03,45)
TP % (sensitivity)	100 (8/8)	100 (8/8)	100 (8/8)	87.5 (7/8)	87.5 (7/8)	87.5 (7/8)	100 (8/8)	100 (8/8)	87.5 (7/8)
TN % (specificity)	75 (3/4)	75 (3/4)	75 (3/4)	75 (3/4)	75 (3/4)	75 (3/4)	75 (3/4)	75 (3/4)	75 (3/4)
FP %	25 (1/4)	25 (1/4)	25 (1/4)	25 (1/4)	25 (1/4)	25 (1/4)	25 (1/4)	25 (1/4)	25 (1/4)
FN %	0 (0/8)	0 (0/8)	0 (0/8)	12.5 (1/8)	12.5 (1/8)	12.5 (1/8)	0 (0/8)	0 (0/8)	12.5 (1/8)
Average Accuracy	0.917	0.917	0.917	0.833	0.833	0.833	0.917	0.917	0.833
Overall Accuracy	0.875	0.875	0.875	0.812	0.812	0.812	0.875	0.875	0.812

Sup. Table 7-3: Model 3 (I_{1u}/IC_{50} or I_2/IC_{50})

Model 3 (I_{1u}/IC_{50} or I_2/IC_{50})	IC_{50}(V)[a]		IC_{50}(ER)[a]		IC_{50}(NSF)[a]	
	ROC Analysis	EMA 2012[e]	ROC Analysis	EMA 2012[e]	ROC Analysis	EMA 2012[e]
Cut-off values	(0.00141,3334)	(0.02,10)	(0.00177,62)	(0.02,10)	(0.00052,94)	(0.02,10)
TP % (sensitivity)	62.5 (5/8)	100 (8/8)	75 (6/8	87.5 (7/8	75 (6/8	100 (8/8)
TN % (specificity)	100 (4/4)	75 (3/4)	75 (3/4	75 (3/4	75 (3/4	75 (3/4)
FP %	0 (0/4)	25 (1/4)	25 (1/4	25 (1/4	25 (1/4	25 (1/4)
FN %	37.5 (3/8)	0 (0/8)	25 (2/8	12.5 (1/8	25 (2/8	0 (0/8)
Average Accuracy	0.75	0.917	0.75	0.833	0.75	0.917
Overall Accuracy	0.812	0.875	0.75	0.812	0.75	0.875

EMA, European Medicines Agency; FDA, US Food and Drug Administration; TN, true negative; TP, true positive; FN, false negative; FP, false positive; IC_{50}, half-maximal inhibitory concentration; $I_{1,}$ the mean steady-state total (free and bound) maximum plasma concentration (C_{max}) of inhibitor following administration of the highest proposed clinical dose; $I_{1,u,}$ unbound I_1; $I_{2,}$ the concentration of inhibitor in the gastrointestinal tract based on highest approved dose dissolved in 250 ml.

[a]: IC_{50}(V), IC_{50}(ER), and IC_{50}(NSF) were used in the static models for DDI prediction, respectively. IC_{50}(V), IC_{50}(ER), and IC_{50}(NSF) was determined as described in the *Materials and Methods*.

[b]: The cut-off value was obtained from FDA final DDI guidance (FDA, 2020)

[c]: The cut-off value was obtained from Ellens et al (2013).

[d]: The cut-off value was obtained from (Agarwal et al., 2013)

[e]: The cut-off value was obtained from EMA DDI guidance (EMA, 2012)

Supplemental Table 8. List of false negative and false positive prediction of digoxin clinical DDIs for test set compounds

		Model 1 (I_2/IC_{50})									Model 2 (I_1/IC_{50} or I_2/IC_{50})									Model 3 (I_{1u}/IC_{50} or I_2/IC_{50})					
		$IC_{50}(V)$[a]			$IC_{50}(ER)$[a]			$IC_{50}(NSF)$[a]			$IC_{50}(V)$[a]			$IC_{50}(ER)$[a]			$IC_{50}(NSF)$[a]			$IC_{50}(V)$[a]		$IC_{50}(ER)$[a]		$IC_{50}(NSF)$[a]	
		ROC Analysis	FDA 2000[b]	Ellens et al. 2012[c]	ROC Analysis	FDA 2000[b]	Ellens et al. 2012[c]	ROC Analysis	FDA 2000[b]	Ellens et al. 2012[c]	ROC Analysis	Agarwal et al. 2012[d]	Ellens et al. 2012[c]	ROC Analysis	Agarwal et al. 2012[d]	Ellens et al. 2012[c]	ROC Analysis	Agarwal et al. 2012[d]	Ellens et al. 2012[c]	ROC Analysis	EMA 2012[e]	ROC Analysis	EMA 2012[e]	ROC Analysis	EMA 2012[e]
	cut-off values	25.9	10	45	13.7	10	45	9.3	10	45	(0.032,40)	(0.1,10)	(0.03,45)	(0.081,26.7)	(0.1,10)	(0.03,45)	(0.026,10)	(0.1,10)	(0.03,45)	(0.00141,3334)	(0.02,10)	(0.00177,62)	(0.02,10)	(0.00052,94)	(0.02,10)
False negative	Cobicistat				√	√	√			√				√	√	√						√	√		
	Tolvaptan			√			√												√	√		√		√	
	Vorapaxar			√																√					
	Neratinib																			√				√	
False positive	Lurasidone	√	√	√	√	√	√	√	√	√	√	√	√	√	√	√	√	√	√		√	√	√	√	√

[a]: $IC_{50}(V)$, $IC_{50}(ER)$, and $IC_{50}(NSF)$ were used in the static models for DDI prediction, respectively. $IC_{50}(V)$, $IC_{50}(ER)$, and $IC_{50}(NSF)$ was determined as described in the *Materials and Methods*.
[b]: The cut-off value was obtained from FDA final DDI guidance (FDA, 2020)
[c]: The cut-off value was obtained from Ellens et al (2013).
[d]: The cut-off value was obtained from (Agarwal et al., 2013)
[e]: The cut-off value was obtained from EMA DDI guidance (EMA, 2012)

Supplemental Table 9. List of false negative and false positive prediction of clinical DDIs using dabigatran etexilate and fexofenadine as P-gp probe drugs

		Model 1 (I_2/IC_{50})								
		IC_{50}(V)[a]			IC_{50}(ER)[a]			IC_{50}(NSF)[a]		
		ROC Analysis	FDA 2000[b]	Ellens et al. 2013[c]	ROC Analysis	FDA 2000[b]	Ellens et al. 2013[c]	ROC Analysis	FDA 2000[b]	Ellens et al. 2013[c]
	cut-off	25.9	10	45	13.7	10	45	9.3	10	45
False negative	Clarithromycin 500mg, BCS III, Dabigatran 150mg						√			√
	Clarithromycin 500mg, BCS III, Dabigatran 0.375mg						√			√
	Clopidogrel 300mg, BCS II, Dabigatran 150mg BID	√		√			√			√
	Cobicistat, BCS II, Dabigatran 150mg				√	√	√			√
	Dronedarone, BCS II, Dabigatran 150mg									√
	Quinidine, BCS I, Dabigatran 150mg									√
	ticagrelor 90mg, BCS IV, Dabigatran 150mg	√		√						
	Verapamil 120mg, BCSII, Dabigatran 150mg									√
	Alogliptin, BCS III, fexofenadine 80mg	√	√	√	√		√			√
	Itraconazole, BCS II, Fexofenadine 180mg									√
	Paroxetine, BCS I, Fexofenadine 60mg			√						√
	Quinidine, BCS I, Fexofenadine 25mg									√
False positive	Atorvastatin, BCS II, Dabigatran 150mg BID	√		√						
	Clopidogrel 75mg, BCS II, Dabigatran 150mg							√		
	Ritonavir, BCS IV, Dabigatran 150mg	√	√	√	√	√		√	√	
	Sertraline, BCS I, Fexofenadine 50mg			√	√	√	√	√	√	

[a]: IC_{50}(V), IC_{50}(ER), and IC_{50}(NSF) were used in the static models for DDI prediction, respectively. IC_{50}(V), IC_{50}(ER), and IC_{50}(NSF) was determined as described in the *Materials and Methods*.
[b]: The cut-off value was obtained from FDA final DDI guidance (FDA, 2020)
[c]: The cut-off value was obtained from Ellens et al (2013).

References in supplemental tables

Agarwal S, Arya V, and Zhang L (2013) Review of P-gp inhibition data in recently approved new drug applications: utility of the proposed [I(1)]/IC(50) and [I(2)]/IC(50) criteria in the P-gp decision tree. *J Clin Pharmacol* **53:**228-233.

Benet LZ (2013) The role of BCS (biopharmaceutics classification system) and BDDCS (biopharmaceutics drug disposition classification system) in drug development. *J Pharm Sci* **102:**34-42.

Bentz J, O'Connor MP, Bednarczyk D, Coleman J, Lee C, Palm J, Pak YA, Perloff ES, Reyner E, Balimane P, Brannstrom M, Chu X, Funk C, Guo A, Hanna I, Heredi-Szabo K, Hillgren K, Li L, Hollnack-Pusch E, Jamei M, Lin X, Mason AK, Neuhoff S, Patel A, Podila L, Plise E, Rajaraman G, Salphati L, Sands E, Taub ME, Taur JS, Weitz D, Wortelboer HM, Xia CQ, Xiao G, Yabut J, Yamagata T, Zhang L, and Ellens H (2013) Variability in P-glycoprotein inhibitory potency

(IC(5)(0)) using various in vitro experimental systems: implications for universal digoxin drug-drug interaction risk assessment decision criteria. *Drug Metab Dispos* **41:**1347-1366.

Ellens H, Deng S, Coleman J, Bentz J, Taub ME, Ragueneau-Majlessi I, Chung SP, Heredi-Szabo K, Neuhoff S, Palm J, Balimane P, Zhang L, Jamei M, Hanna I, O'Connor M, Bednarczyk D, Forsgard M, Chu X, Funk C, Guo A, Hillgren KM, Li L, Pak AY, Perloff ES, Rajaraman G, Salphati L, Taur JS, Weitz D, Wortelboer HM, Xia CQ, Xiao G, Yamagata T, and Lee CA (2013) Application of receiver operating characteristic analysis to refine the prediction of potential digoxin drug interactions. *Drug Metab Dispos* **41:**1367-1374.

EMA C (2012) Guideline on the investigation of drug interactions European Medicines Agency. *https://wwwemaeuropaeu/documents/scientific-guideline/guideline-investigation-drug-interactions_enpdf*.

FDA C (2020) In Vitro Drug Interaction Studies — Cytochrome P450 Enzyme- and Transporter-Mediated Drug Interactions Guidance for Industry, in: *Center for Drug Evaluation and Research (CDER)*.

Fekete Z, Rajnai Z, Nagy T, Jakab KT, Kurunczi A, Gemes K, Heredi-Szabo K, Fulop F, Toth GK, Czerwinski M, Loewen G, and Krajcsi P (2015) Membrane Assays to Characterize Interaction of Drugs with ABCB1. *J Membr Biol* **248:**967-977.

Fujita Y, Noguchi K, Suzuki T, Katayama K, and Sugimoto Y (2013) Biochemical interaction of anti-HCV telaprevir with the ABC transporters P-glycoprotein and breast cancer resistance protein. *BMC Res Notes* **6:**445.

Heredi-Szabo K, Palm JE, Andersson TB, Pal A, Mehn D, Fekete Z, Beery E, Jakab KT, Jani M, and Krajcsi P (2013) A P-gp vesicular transport inhibition assay - optimization and validation for drug-drug interaction testing. *Eur J Pharm Sci* **49:**773-781.

Horio M, Gottesman MM, and Pastan I (1988) ATP-dependent transport of vinblastine in vesicles from human multidrug-resistant cells. *Proc Natl Acad Sci U S A* **85:**3580-3584.

Lempers VJ, van den Heuvel JJ, Russel FG, Aarnoutse RE, Burger DM, Bruggemann RJ, and Koenderink JB (2016) Inhibitory Potential of Antifungal Drugs on ATP-Binding Cassette Transporters P-Glycoprotein, MRP1 to MRP5, BCRP, and BSEP. *Antimicrob Agents Chemother* **60:**3372-3379.

Papich MG and Martinez MN (2015) Applying Biopharmaceutical Classification System (BCS) Criteria to Predict Oral Absorption of Drugs in Dogs: Challenges and Pitfalls. *AAPS J* **17:**948-964.

Safar Z, Jani M, Makai I, Fekete Z, Bui A, Molnar E, Padar P, Pratt JR, Kis E, Beery E, and Krajcsi P (2018) Correlation Analysis of Potential Breast Cancer Resistance Protein Probes in Different Monolayer Systems. *J Pharm Sci* **107:**2742-2747.

Schaefer M, Roots I, and Gerloff T (2006) In-vitro transport characteristics discriminate wild-type ABCB1 (MDR1) from ALA893SER and ALA893THR polymorphisms. *Pharmacogenet Genomics* **16:**855-861.

Vermeer LM, Isringhausen CD, Ogilvie BW, and Buckley DB (2016) Evaluation of Ketoconazole and Its Alternative Clinical CYP3A4/5 Inhibitors as Inhibitors of Drug Transporters: The In Vitro Effects of Ketoconazole, Ritonavir, Clarithromycin, and Itraconazole on 13 Clinically-Relevant Drug Transporters. *Drug Metab Dispos* **44:**453-459.

Wu CY and Benet LZ (2005) Predicting drug disposition via application of BCS: transport/absorption/ elimination interplay and development of a biopharmaceutics drug disposition classification system. *Pharm Res* **22:**11-23.