Xenobiotica
the fate of foreign compounds in biological systems

ISSN: 0049-8254 (Print) 1366-5928 (Online) Journal homepage: https://www.tandfonline.com/loi/ixen20

Comparison of methods for the prediction of human clearance from hepatocyte intrinsic clearance for a set of reference compounds and an external evaluation set

Tetsuo Yamagata, Ugo Zanelli, Dieter Gallemann, Dominique Perrin, Hugues Dolgos & Carl Petersson

To cite this article: Tetsuo Yamagata, Ugo Zanelli, Dieter Gallemann, Dominique Perrin, Hugues Dolgos & Carl Petersson (2017) Comparison of methods for the prediction of human clearance from hepatocyte intrinsic clearance for a set of reference compounds and an external evaluation set, Xenobiotica, 47:9, 741-751, DOI: 10.1080/00498254.2016.1222639

To link to this article: https://doi.org/10.1080/00498254.2016.1222639

Published online: 25 Aug 2016.

http://tandfonline.com/ixen
ISSN: 0049-8254 (print), 1366-5928 (electronic)

Xenobiotica

Xenobiotica, 2017; 47(9): 741–751
 DOI: 10.1080/00498254.2016.1222639

Taylor & Francis
Taylor & Francis Group

RESEARCH ARTICLE

Comparison of methods for the prediction of human clearance from hepatocyte intrinsic clearance for a set of reference compounds and an external evaluation set

Tetsuo Yamagata[1], Ugo Zanelli[2], Dieter Gallemann[1], Dominique Perrin[2], Hugues Dolgos[2], and Carl Petersson[2]

[1]*Global Early Development/Quantitative Pharmacology and Drug Disposition (QPD), Merck KGaA, Grafing, Germany and* [2]*Global Early Development/Quantitative Pharmacology and Drug Disposition (QPD), Merck KGaA, Darmstadt, Germany*

Abstract

1. We compared direct scaling, regression model equation and the so-called ''Poulin et al.'' methods to scale clearance (CL) from *in vitro* intrinsic clearance (CL_{int}) measured in human hepatocytes using two sets of compounds. One reference set comprised of 20 compounds with known elimination pathways and one external evaluation set based on 17 compounds development in Merck (MS).
2. A 90% prospective confidence interval was calculated using the reference set. This interval was found relevant for the regression equation method. The three outliers identified were justified on the basis of their elimination mechanism.
3. The direct scaling method showed a systematic underestimation of clearance in both the reference and evaluation sets. The ''Poulin et al.'' and the regression equation methods showed no obvious bias in either the reference or evaluation sets.
4. The regression model equation was slightly superior to the ''Poulin et al.'' method in the reference set and showed a better absolute average fold error (AAFE) of value 1.3 compared to 1.6. A larger difference was observed in the evaluation set were the regression method and ''Poulin et al.'' resulted in an AAFE of 1.7 and 2.6, respectively (removing the three compounds with known issues mentioned above). A similar pattern was observed for the correlation coefficient. Based on these data we suggest the regression equation method combined with a prospective confidence interval as the first choice for the extrapolation of human *in vivo* hepatic metabolic clearance from *in vitro* systems.

Keywords

Liver, metabolism, scaling, well-stirred model, *in vitro–in vivo* extrapolation, Poulin et al. method, regression equation

History

Received 28 June 2016
Revised 5 August 2016
Accepted 7 August 2016
Published online 24 August 2016

Introduction

The majority of marketed drugs as well as drugs under development has been shown to be eliminated mainly via hepatic metabolism (Williams et al., 2004). Consequently, there is a need for methods that are able to predict human clearance of compounds eliminated via this clearance pathway in drug discovery as well as in preclinical development. The prediction of human clearance is a central part of human dose predictions which was recently highlighted in several cooperate drug discovery strategies (Cook et al., 2014; Dolgos et al., 2016). Isolated hepatocytes provide an intact cellular system containing a full complement of drug metabolizing enzymes, transporters and cofactors, making them ideal for studying rates of drug metabolism corresponding to the industry needs (Brandon et al., 2003; Donato & Castell, 2003; Hewitt et al., 2007). The utility of these *in vitro* intrinsic clearance (CL_{int}) assays for rank ordering of compounds based on their intrinsic metabolic stability and for the extrapolation to predictions of metabolic clearance in preclinical species and humans has been widely demonstrated (Hallifax et al., 2008; Houston, 1994; McGinnity et al., 2004). The main limitation of the approach is that metabolism rather than transport has to be the rate limiting step. Hence, clearance of renal and biliary cleared compounds will be underestimated. Similarly, some metabolically cleared compounds for which the elimination is limited by uptake rate rather than metabolism have been shown to be heavily underestimated by standard set ups (Soars et al., 2007). Adaptations of the standard method were suggested to address the clearance of this type of molecules (Nordell et al., 2013).

Extrapolation of hepatocyte-derived CL_{int}s (the so-called direct scaling) results in a systematic underestimation of the *in vivo* value, despite incorporation of established physiological scaling factors (SFs) and the unbound fractions in both blood and *in vitro* matrix (Brown et al., 2007; Ito & Houston, 2004; Obach, 1999; Poulin et al., 2012a; Riley et al., 2005; Stringer et al., 2008). There are a number of plausible

Address for Correspondence: C. Petersson, Global Early Development/Quantitative Pharmacology and Drug Disposition (QPD), Frankfurter Str. 250 I Postcode D39/001, 64293 Darmstadt, Germany. E-mail: carl.petersson@merckgroup.com

explanations for this observation. For example, it is known that the *in vitro* incubation conditions employed can greatly influence the rate of drug metabolism (Grime & Riley, 2006). Empirical approaches have been successfully applied to correct for the underpredictions observed across a range of drugs (Ito & Houston, 2005; Riley et al., 2005; Sohlenius-Sternbeck et al., 2012). An attractive alternative method has recently been reported to be superior to the empirical methods discussed above. This method invokes further physiological factors, which accounts for the pH differences and the protein-facilitated uptake due to potential ionic interactions between the protein-bound drug complex and cell surface of the hepatocytes (Poulin et al., 2012a,b; Poulin & Haddad, 2013). This approach appears especially attractive as it provides an explanation for the observed underestimation, furthermore it does not rely on a number of references to provide adequate correction factors. In this work we compare the performance of the empirical so-called ''regression equation'' method to direct scaling and to the ''Poulin et al.'' method using a reference set, consisting of compounds with well-defined human clearance pathways, and an external evaluation set. The external evaluation set was composed of all small molecules (namely 17) developed within Merck over the 15 years for which a human clearance value was available and an CL_{int} could be determined *in vitro*. This set was critical to method comparison as the selection of the reference compounds and the associated literature data might have distorted correlations and therefore make method comparisons difficult. We used the external evaluation set to explore the validity by the two-sided prospective prediction interval defined using the reference set suggested by Sohlenius-Sternbeck et al. (2012) as an indicator of the accuracy in a truly external evaluation set in a human system.

Materials and methods

Chemicals

Carvedilol, chlorpromazine, desipramine, diltiazem, etodolac, fenoprofen, ketoprofen, metoprolol, midazolam, naloxone, omeprazole, pindolol, prednisolone, propranolol, quinidine, sildenafil, tolbutamide, verapamil and zidovudine were purchased from Sigma-Aldrich Corporation, LLC. (St. Louis, MO). Diazepam, and all Merck developmental compounds (asimadoline, ceralifimod, eniporide, mitoxantrone, pimasertib, pruvanserin and 11 Merck clinical candidates) were obtained from Merck (Darmstadt, Germany). All other chemicals were of analytical grade and obtained from commercial suppliers.

In vitro incubation in human hepatocyte suspension

Cryopreserved human hepatocytes (pools of 100 donors) were obtained from XenoTech, LLC. (Kansas City, KS) and human sera was obtained from Merck internal blood donation program. Each vial of cryopreserved hepatocytes was thawed according to the manufacturer's instructions. Hepatocyte viabilities were determined by trypan blue exclusion after thawing. In a dedicated experiment viability was also measured at 90, 120, 180 and 360 min. All hepatocyte incubations were performed in 96-well plate (Multiwell plate, round bottom, VWR International GmbH, Darmstadt, Germany). Test compounds (reference compounds and Merck compounds) were incubated at a concentration of 0.25 μM with the hepatocytes (1 or 2×10^6 cells/mL) that were suspended in Krebs-Henseleit buffer. In order to avoid possible stickiness effect and improve viability during the incubation, 5% human serum was present in the incubations with all the test compounds, except for MS11, with which 1% serum was present as its plasma protein binding was extremely high. The incubations were carried out in duplicate (at least) at 37 °C under an atmosphere of 5% CO_2 and the plates were shaken gently. In order to minimize an evaporation of the incubation medium, the incubation plates were placed in wet chambers filled with paper towels during the incubation and the edge wells of the incubation plates were not used. Aliquots were taken at 10 time points (e.g. 0, 0.08, 0.25, 0.5, 1, 2, 3, 3.7, 4.3 and 5 h) and the reactions were stopped by the addition of ice-cold acetonitrile/water (1:1). After the centrifugation of the samples, the supernatants were analyzed by LC-MS/MS.

Determination of CL_{int}

The disappearance of the compounds was expressed as % remaining relative to the concentration measured at the 0 min incubation and the logarithm % remaining values were plotted against time. The curves in the initial linear range were fitted to the first-order elimination equation (Equation (1)) using GraphPad Prism (version 5.04, GraphPad Software Inc., San Diego, CA):

$$R_t = R_0 \times e^{-k_{el} \times t} \quad (1)$$

where R_t is the percentage remaining at any time, R_0 is the percentage remaining at 0 min (= 100%) and k_{el} is the first-order elimination constant. The CL_{int} normalized by the number of hepatocytes in the incubation was calculated as follows:

$$CL_{int} = \frac{k_{el} \times V}{n} \quad (2)$$

where V is the volume of the incubation and n is the number of hepatocytes.

Calculation of fraction unbound in incubation with human hepatocyte in the presence of serum

The fraction unbound to hepatocytes in incubation (fu_{hep}) can be predicted, according to Kilford et al. (2008):

$$fu_{hep} = \frac{1}{1 + 125 \times V_R \times 10^{0.072 \times \log P/D^2 + 0.067 \times \log P/D - 1.126}} \quad (3)$$

where V_R is the ratio between the cell volume and the incubation volume and it has a value of $0.005/10^6$ cells/mL. For this calculation, log$P$ applied for bases and log$D^{7.4}$ for acids and neutrals.

As human serum was present in the medium protein binding was taken into account when calculating the free fraction in the incubation. The overall fraction unbound in the incubations ($fu_{p,hep}$) was estimated using unbound fraction in

human plasma, assuming no difference in binding in human plasma and serum:

$$fu_{\mathrm{p,hep}} = \frac{1}{\left[\left(1/fu_{\mathrm{p}} - 1\right)/d + 1/fu_{\mathrm{hep}}\right]} \quad (4)$$

where fu_{p} is unbound fraction in human plasma and d is the dilution factor of serum (e.g. $d = 20$ in 5% serum).

Plasma protein binding

The unbound fraction in plasma (fu_{p}) values for the reference compounds were obtained from Sohlenius-Sternbeck et al. (2012) except for zidovudine that was obtained from Luzier & Morse (1993).

The fu_{p} for the Merck compounds was determined *in vitro* using a concentration of 0.1–100 μM by equilibrium dialysis. Human plasma was dialyzed against phosphate buffer at pH 7.4 for 24 h. Following equilibration at 37 °C, the concentrations of the compound in both plasma and buffer compartments were measured by either LC-MS/MS, or liquid scintillation counting (LSC) for nonradio-labeled or radio-labeled materials, respectively. The fu_{p} was determined by dividing the concentration in the dialysate by the one in the plasma. If saturation was observed the fu_{p} observed at the lowest concentration was utilized. The fu_{p} of MS2 and MS5 was however assessed by ultrafiltration with Multiscreen Filter Plate System with Ultracel-10 Membrane (MerckMillipore, Darmstadt, Germany), after incubation of the compound 5 μM at 37 °C for 30 min, in human plasma serum diluted 1:2 and 1:5. The concentration in the ultrafiltrate, as the free concentration, was measured by LC-MS/MS.

Plasma protein-binding identification

To identify the main plasma protein-binding partner, equilibrium dialysis in the purified plasma protein solutions, spiked with 600–750 μM human serum albumin (HSA) or 18–22 μM α1-acid glycoprotein (AAG) solution, were carried out and the unbound fraction in each plasma protein was determined. The plasma protein which showed the lower unbound fraction was defined as the main binding partner. When no experimental value was available, binding to AAG was assumed for basic compounds, whereas binding to HSA was assumed for acidic and neutral compounds as described in the literature (Poulin & Haddad, 2013).

Blood-to-plasma partition ratio (C_B/C_P)

Data of blood-to-plasma partition ratios (C_B/C_P) of reference compounds was obtained from the literature (Luzier & Morse, 1993; Sohlenius-Sternbeck et al., 2012). For the Merck compounds, the C_B/C_P ratios were experimentally determined. Briefly, the compound was spiked into human fresh blood and was incubated at 37 °C for 30 min. The amount of the compound in plasma was measured by LC-MS/MS or LSC. The C_B/C_P was calculated by dividing the initial concentration in blood by the concentration in plasma. When no experimental value was available, a value of 1 was assumed for neutral and basic compounds and 0.55 for acidic compounds.

IVIVE methods

In this study, three *in vitro-in vivo* extrapolation (IVIVE) methods were evaluated, (1) ''direct scaling method,'' (2) ''Poulin et al. method'' and (3) ''regression equation method.'' The corresponding equations are described in Table 1.

The physiological SF utilized were identical for all methods, the numbers utilized were obtained from SimCYP program V13.1 module human (Certara Inc., Princeton, NJ). The hepatocellularity and the liver weight utilized were 118×10^6 cells/g liver and 1718 g liver/75 kg body weight, respectively. The utilized human liver blood flow was 19.3 mL/min/kg body weight (SimCYP program V13.1)

Table 1. Overview of applied extrapolation methods.

Methods	Model Equations
Direct scaling method	$\mathrm{Pred}.L_H = \frac{Q_H \times \left[\frac{CL_{\mathrm{int}}}{fu_{hep,p}} \times SF \times \frac{fu_p}{C_B\ /\ C_P}\right]}{Q_H + \left[\frac{CL_{\mathrm{int}}}{fu_{hep,p}} \times SF \times \frac{fu_p}{C_B\ /\ C_P}\right]}$
Poulin et al. method	$\mathrm{Pred}.CL_H = \frac{Q_H \times \left[\frac{CL_{\mathrm{int}}}{fu_{hep,p}} \times SF \times \frac{fu_{\mathrm{liver}}}{C_B\ /\ C_P}\right]}{Q_H + \left[\frac{CL_{\mathrm{int}}}{fu_{hep,p}} \times SF \times \frac{fu_{\mathrm{liver}}}{C_B\ /\ C_P}\right]}$
Regression equation method	$\mathrm{Pred}.CL_H = \frac{Q_H \times \left[\frac{CL_{\mathrm{int,corr}}}{fu_{hep,p}} \times SF \times \frac{fu_p}{C_B\ /\ C_P}\right]}{Q_H + \left[\frac{CL_{\mathrm{int,corr}}}{fu_{hep,p}} \times SF \times \frac{fu_p}{C_B\ /\ C_P}\right]}$ where $\log\left[\left(\frac{CL_{\mathrm{int,corr}}}{fu_{hep,p}} \times SF \times \frac{fu_p}{C_B\ /\ C_P}\right)\right] = \mathrm{Slope} \times \log\left[\left(\frac{CL_{\mathrm{int}}}{fu_{hep,p}} \times SF \times \frac{fu_p}{C_B\ /\ C_P}\right)\right] + \mathrm{Intercept}$

$\mathrm{Pred}.CL_H$: predicted hepatic clearance; CL_{int}: intrinsic clearance; $fu_{\mathrm{hep,p}}$: unbound fraction in *in vitro* hepatocytes incubation containing serum; Q_H: liver blood flow; fu_{p}: unbound fraction in plasma; C_B/C_P: blood-to-plasma partition ratio; fu_{liver}: unbound fraction in liver considering the protein-facilitated uptake and pH gradient, which is well detailed in the original manuscripts (Poulin et al., 2012a,b; Poulin & Haddad, 2013); SF: physiological scaling factor, i.e. (118×10^6 cells/g liver) × (1718 g liver/75 kg body weight); $CL_{\mathrm{int,corr}}$: intrinsic clearance corrected using reference compounds.

Briefly, in direct scaling method, the *in vitro* raw CL_{int} values are scaled to a hepatic intrinsic clearance ($CL_{int,H}$) using physiological SF, $fu_{p,hep}$, fu_p and C_B/C_P. The $CL_{int,H}$ is applied to the well-stirred model to predict the hepatic *in vivo* blood clearance.

The ''Poulin et al.'' method further considers (1) a difference of drug ionization between the plasma and intracellular water in liver and (2) a possible delivery of additional free drug to the hepatocytes via a protein-facilitated uptake, which is estimated by correcting the apparent fu_p values of each drug for the differences of HSA concentration between the plasma and liver, and thus, fu_{liver} is used as the correction factor. The utility of fu_{liver} is applied only for a drug bound mainly to HSA because for a drug principally bound to AAG, the apparent fu_p value was used instead. The calculation of fu_{liver} is described in the original literature (Poulin et al., 2012a,b; Poulin & Haddad, 2013).

The regression equation method corrects for the systematic bias by applying empirical correction derived from regression analysis between the *in vitro* estimated $CL_{int,H}$ and *in vivo* $CL_{int,H}$, back-calculated from the observed *in vivo* blood CL based on the well stirred model, within a defined set of reference compounds (Sohlenius-Sternbeck et al., 2012).

Human pharmacokinetic data

The clearances of the reference compounds were collected from literature sources, these are summarized in Table 2. The external evaluation set was composed of all small molecules (namely 17) developed within Merck over 15 years for which a human clearance value was available and CL_{int} value could be determined *in vitro* are summarized in Table 3. Wherever possible literature references are given. The clearance of compounds for which no bioavailability data (5 of 17) available was estimated assuming no gastrointestinal loss ($F_{gut} \times F_{abs} = 1$).

Statistics

Performance evaluation

From the IVIVE plots, a measure of the bias, expressed as the average fold error (AFE), that is, the geometric mean fold error, was calculated as follows (Tang et al., 2007):

$$AFE = 10^{\frac{1}{N} \times \sum^{Log\ Y_{Pred}/Y}} \quad (5)$$

The average absolute fold error (AAFE), a measure of the precision, was calculated using the following equation (Tang et al., 2007):

$$AAFE = 10^{\frac{1}{N} \times \sum^{|Log Y_{Pred}/Y|}} \quad (6)$$

Prospective prediction interval

The prospective prediction interval was calculated as described by Sohlenius-Sternbeck et al., (2012) and

Table 2. Reference compounds with known hepatic elimination used for the construction of regression line, their ion class, lipophilicity, p*K*a, main binding partner, *C*B/*C*P, observed CL_B and predicted clearance using direct scaling, the regression and Poulin et al. methods.

Drug	Ion class	Log *P*/*D*[a]	p*Ka*[b]	Main binding protein[b]	fu_p[a]	CB/CP[a]	Raw *in vitro* CL_{int} (μL/min/10^6cells)	Observed CL_B[d] (mL/min/kg)	Direct scaling (mL/min/kg)	Regression equation (mL/min/kg)	Poulin et al. (mL/min/kg)
Etodolac	Acid	0.5	4.7	HSA	0.008	0.60	1.7	1.3	0.43	1.5	2.3
Fenoprofen	Acid	0.58	4.5	HSA	0.002	0.55	3.2	1.7	0.79	2.6	4.0
Ketoprofen	Acid	−0.14	4.45	HSA	0.004	0.55	0.63	2.0	0.18	0.66	0.98
Tolbutamide	Acid	0.40	5.27	HSA	0.021	0.68	0.56	0.35	0.16	0.60	0.84
Carvedilol	Base	4.31	8.1	HSA	0.009	0.70	13	9.9	3.0	7.7	16
Chlorpromazine	Base	5.43	9.7	AAG	0.016	0.78	18	9.8	9.3	15	14
Desipramine	Base	4.00	10.3	AAG	0.16	0.96	6.2	11	5.1	11	9.0
Diltiazem	Base	2.79	7.7	AAG	0.25	0.98	6.9	13	5.0	10	7.7
Metoprolol	Base	1.75	9.7	AAG	0.883	1.07	1.6	12	3.4	8.2	4.0
Naloxon	Base	2.28	7.9	AAG	0.722	1.22	28	18	14	17	15
Pindolol	Base	1.86	8.8	AAG	0.84	0.69	1.1	5.2	3.6	8.6	4.4
Propranolol	Base	3.42	9.5	AAG	0.205	0.86	13	15	8.5	14	12
Quinidine	Base	3.55	10, 5.4	AAG	0.18	0.90	2.0	4.9	1.8	5.2	3.7
Sildenafil	Base	2.8	7.6	AAG	0.066	0.63	11	8.0	4.5	9.8	7.6
Verapamil	Base	3.96	8.5	AAG	0.129	0.77	20	15	10	15	14
Diazepam	Neutral	2.81		HSA	0.013	0.78	0.80	0.46	0.18	0.66	2.1
Midazolam	Neutral	3.4		HSA	0.023	0.67	9.1	6.6	2.7	7.0	13
Omeprazole	Neutral	2.11		HSA[e]	0.18	0.60	3.6	14[c]	3.6	8.5	10
Prednisolone	Neutral	1.6		HSA	0.483	1.00	3.6	5.8	4.2	9.5	7.5
Zidovudine	Neutral	0.05[f]		HSA	0.77[g]	0.86[g]	3.5	12.4	6.2	12	8.0

fu_p: unbound fraction in plasma; C_B/C_P: blood-to-plasma partition ratio; CL_{int}: intrinsic clearance; CL_B: blood clearance; HSA: human serum albumin, AAG: alfa 1-acid glycoprotein.

[a]Sohlenius-Sternbeck et al., 2012, unless other is indicated.

[b]Poulin & Haddad, 2013.

[c]Sohlenius-Sternbeck et al., 2012.

[d]Poulin & Haddad, 2013; the mean value was calculated in case two values were available.

[e]Assumption based on the ion class.

[f]Hansch et al., 1995.

[g]Luzier & Morse, 1993.

Table 3. Merck compounds, their main elimination pathways, ion class, lipophilicity, pKa, main binding partner, CB/CP, observed CL_B and predicted clearance using direct scaling, the regression and Poulin et al. methods.

Drug	Ion class	Log *P/D*	p*K*a	Main binding protein	fu_p[b]	CB/CP	Raw *in vitro* CL_{int} (μL/min/10^6 cells)	Observed CL_B[g] (mL/min/kg)	Direct scaling (mL/min/kg)	Regression equation (mL/min/kg)	Poulin et al. (mL/min/kg)	Comment
MS1	Neutral	4.3		HSA[a]	0.020	0.70	0.32	0.36[e]	0.17	0.66	1.7	
MS2	Base	4.5	8.3	AAG	0.038[c]	0.75	1.2	1.4	1.2	3.7	1.2	
MS3	Base	4.5	8.4	AAG[a]	0.20	1.1	3.0	10[e]	6.2	12	9.0	
MS4	Base	3.6	9.5	HSA	0.020	0.97	1.6	3.1	0.42	1.5	4.0	
MS5	Neutral	2.1		HSA[a]	0.085[c]	1.2	3.0	2.9[e]	1.0	3.2	5.1	
MS6	Base	2.7	7.6	HSA	0.115	0.74	0.88	1.6	0.66	2.2	3.7	
MS7	Base	4.7	7.7	HSA	0.022	0.98	1.7	1.6[e]	1.1	3.4	9.7	
MS8	Neutral	−3.5		HSA[a]	0.80	1.2	0.29	3.6	0.72	2.4	0.88	
MS9	Neutral	3.2		HSA	0.060	1.3	6.2	5.5	1.9	5.4	8.8	
MS10	Neutral	1.1		HSA	0.55	1.0	4.1	16	5.2	11	7.5	Extrahepatic metabolism[i]
MS11	Base	4.8	7.6	HSA	0.000020	0.60	3.6	1.4	0.16	0.61	3.4	
Asimadoline	Base	4.2	7.9	AAG[a]	0.080	0.70	5.3	1.2	5.6	11	8.6	Enterohepatic recirculation[j]
Ceralifimod	Base	4.3	9.8	HSA	0.030	0.93	0.92	0.36[e,k]	0.49	1.7	5.9	
Eniporide	Neutral	0.58		HSA[a]	0.28	0.90	3.2	4.7[f]	2.9	7.3	6.6	
Mitoxantrone	Base	2.8	8.2,7.5	AAG[a]	0.22	1.0[a]	0.83	4.8[g,h]	0.81	2.7	1.9	
Pimasertib	Neutral	2.4		HSA	0.12	1.1	2.8	9.6[e]	1.2	3.8	5.1	Extrahepatic metabolism[i]
Pruvanserin	Neutral	2.9		HSA[a]	0.051	0.70	4.2	6.6[e]	1.9	5.4	2.2	

fu_p: unbound fraction in plasma; C_B/C_P: blood-to-plasma partition ratio; CL_{int}: intrinsic clearance; CL_B: blood clearance; HSA: human serum albumin; AAG: alfa 1-acid glycoprotein.

[a]Assumption based on the ion class.

[b]The free fraction was estimated via equilibrium dialysis in a regulatory setting unless otherwise stated.

[c]Ultrafiltration was used to estimate the free fraction.

[d]Data acquired from internal reports unless stated.

[e]Clearance was derived based on the assumption that the gastrointestinal loss was negligible.

[f]Richard et al., 1992.

[g]Kovar et al., 2001.

[h]The value reflects only the hepatic component, the renal part has been subtracted.

[i]Internal Merck reports.

[j]Reported by Camilleri, 2008.

[k]Krösser et al., 2015.

[l]Ravandi et al., 2015.

Armitage & Berry (1994). The prediction limits/interval were calculated as follows:

$$\mathrm{Log10}(Y) = \mathrm{Intercept} + \mathrm{slope} * \mathrm{Log10}(X) + s*t*\sqrt{1 + \frac{1}{N} + \frac{(X - \bar{X})^2}{Exx}} \quad (7)$$

where Y is the derived *in vivo* CL_{int}, X is the scaled $CL_{int,H}$, s, t, N, $\bar{X}$, Exx are all derived by the regression.

s = standard deviation (SD) about regression line;

t = critical Student t-value with N-2 degrees of freedom (df);

N = number of observations forming the regression;

$\bar{X}$ = mean of log10 (scaled CL_{int});

Exx = corrected sum of squares of log10 (scaled CL_{int}).

Results

Cell viability

Hepatocyte viability determined by trypan blue exclusion after thawing was >75%. In the absence of serum, cell viability was maintained up to 1.5 h and rapidly declined below 50% after 3 h. The viability was maintained beyond 3 h in the presence of 5% serum.

Reference compounds

The reference compounds used to derive the human linear regression are shown in Table 2. This table also provides information on ion class, fu_p and C_B/C_P values, calculated or measured log*P/D*, blood clearance (CL_B), *in vitro* raw CL_{int}, main binding partner and predicted *in vivo* CL_B using different extrapolation approaches. Reference compounds were selected on the basis of having hepatic metabolic clearance as the principal route of elimination. However, a fraction of the reference compounds had a known renal clearance component which was subtracted from the total clearance in order to study hepatic metabolic elimination exclusively (Poulin & Haddad, 2013, Sohlenius-Sternbeck et al., 2012). This set of compounds represent several ion classes and a wide clearance range. Using the 20 reference compounds, the following regression line below was obtained.

$$\mathrm{Log}\left[\frac{\mathrm{CL}_B \times Q_H}{Q_H - CL_B}\right] = 0.969 \times \log[\mathrm{CL}_{\mathrm{int},H}] + 0.562 \quad (8)$$

The resulting regression line had an R^2 of 0.87 (Figure 1).

Table 4 summarize the outcome of the comparison of the three different extrapolation methods. Similar to earlier studies, it was found that the direct scaling resulted in a significant underestimation of clearance with an AFE of 0.43, whereas the regression equation and the ''Poulin et al.'' method showed no bias. This underestimation resulted in relatively poor accuracy for the direct scaling method with an AAFE of 2.3 and only 35% within twofold. The regression equation method showed slightly better accuracy than the ''Poulin et al.'' method. Marked differences were observed in correlation coefficients (R^2), where the regression equation method resulted in 0.79, whereas the ''Poulin et al.'' method only produced 0.56. These differences resulted in a more narrow perspective confidence interval for the regression equation method compared to the ''Poulin et al.'' method (Figure 2).

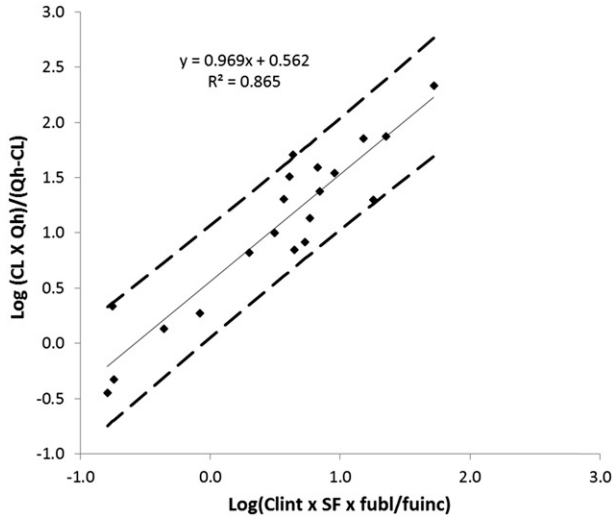

Figure 1. Human hepatocyte regression line established with unbound *in vitro* CL_{int} correlated with CL_{int} *in vivo* corrected with fu_b.

Evaluation set of Merck compounds

The data for the evaluation set, composed of historical Merck clinical candidates that have been in development, are summarized in Table 3. The set contains all compounds for which human clearance could be derived and a CL_{int} could be measured in hepatocytes developed during the last 15 years within Merck. The table also summarizes the variables used in the extrapolations and the extrapolation results for the three examined methods similar to Table 2. In this set, any known renal clearance was subtracted in a way similar to the reference set. No correction for other nonhepatic clearance pathways was made but when such pathways were evident, this is indicated in the table.

A significant underestimation was observed using direct scaling with an AFE of 0.43 similar to the reference set (Table 4). No bias was observed for the regression equation method but the ''Poulin et al.'' method did overestimate the clearance, AFE with an of 1.5. This difference was largely due to one single compound (ceralifimod), which was underestimated by more than 10-fold, hence a systematic bias did not appear as evident. The accuracy (expressed as % within twofold and AAFE) and R^2 values of both the ''Poulin et al.'' and the regression equation methods were lower than for the reference compounds.

The 90% confidence interval calculated using the reference set for the direct scaling method indicated that pimasertib had additional clearance pathway(s). The same interval calculated for the regression equation model additionally indicated that, pimasertib, and MS10 had additional clearance pathway(s). Asimadoline and ceralimod were significantly overestimated and found outside the 90% confidence interval. The ''Poulin et al.'' method did not detect any compound with additional clearance pathways, it did however overestimate the clearance of not only asimadoline and ceralifimod but also MS1 and MS7 to a degree that they were found outside the 90% confidence (Figure 3).

Table 4. Accuracy of extrapolation methods.

	Reference compounds (training set) $N=20$			Prospective compounds (evaluation set) ($N=17$)			Prospective compounds (evaluation set) hepatic elimination ($N=14$)		
Measure	Direct scaling	Regression equation	Poulin et al.	Direct scaling	Regression equation	Poulin et al.	Direct scaling	Regression equation	Poulin et al.
AFE	0.43	1.0	1.1	0.40	1.2	1.4	0.37	1.2	1.5
Absolute average fold error (AAFE)	2.3	1.3	1.6	3.1	1.9	2.6	2.8	1.7	2.6
% within twofold	35	95	90	29	59	41	36	64	43
% outside 90% prospective confidence interval	NA	NA	NA	6	24	24	0	7	21
R^2	0.71	0.79	0.57	0.34	0.40	0.09	0.72	0.75	0.11

Removal of the two compounds with a reported extrahepatic metabolism (pimasertib and MS10) and asimadoline, which has been reported to show a large enterohepatic recirculation, resulted in an increased precision and correlation similar to those observed for the reference compounds for the regression equation and the direct scaling methods (Camilleri, 2008, Scheible et al., in manuscript). By contrast, the ''Poulin et al.'' method still showed a significantly lower $R^2$and decreased accuracy in the evaluation set (Figure 3, Table 4).

Discussion

Several studies have indicated that the application of direct scaling results in an underestimation of clearance. This study confirms these findings and underpins the need for a correction of this systematic error (Brown et al., 2007; Ito & Houston, 2004; Obach, 1999; Poulin et al., 2012a; Riley et al., 2005; Stringer et al., 2008). This work further supports the theory that this underestimation has an assay-dependent component. The observed slope and the intercept in this study differs from what has been reported previously (Riley et al., 2005; Sohlenius-Sternbeck et al., 2010; Sohlenius-Sternbeck et al., 2012). The fact that these studies do not show identical results makes it tempting to speculate that the addition of serum (5%) may affect the slope and intercept. An alternate explanation for the differences in slope and intercept may be lab-to-lab variability in the *in vitro* CL_{int} determinations which has been discussed previously in the literature (Sohlenius-Sternbeck et al., 2012).

The serum addition to the incubation did improve the viability of the hepatocytes in suspension and thereby allowed more data points to be sampled over a longer time. This, combined with measures to minimize evaporation resulted in the quantification of CL_{int} in a range that has been below the detection limit in earlier studies (Stringer et al., 2008). Furthermore, the long incubation time and the careful control of evaporation also allowed sampling of more time points over a wider time interval which may also explain the high correlation coefficient (0.79), an AAFE of 1.3 and 95% of compounds within twofold for the reference set. These values were slightly higher than that reported by Sohlenius-Sternbeck et al. (2010, 2012) (76 and 66% within twofold, respectively) but comparable to Poulin et al (2012a) which reported 88% within twofold. The composition of the datasets may introduce artificially high or low results complicating any method comparison. In addition, the selection of literature values (e.g. fu_b) may contribute to complicate method comparisons. This means that any study comparing the results of different extrapolation methods could be challenged unless all results including *in vivo* were generated within the same laboratory. This degree of data homogeneity may be hard or impossible to obtain. In the current study, we try to address these concerns by using an external evaluation set. This set consisted of all small molecules developed by Merck over the last 15 years for which only one set *in vivo* clearance, blood plasma ratios and fu_b values were available, hence no preselection of the studied compounds were made. The objective of this study was to compare the extrapolation methods of clearance, in order to contribute to this comparison a measurable CL_{int} has to be obtainable. Consequently, all compounds ($N=5$) with no detectable CLint in the hepatocyte incubation were excluded. The evaluation set was different from the reference set as the CLint range was much lower, about almost threefold lower on average (2.5 versus 7.4 μL/min/10^6 cells for the reference and evaluation set, respectively). This is likely to be explained by the practice of optimizing and selecting candidates with low metabolic clearance typical for Merck as well as the majority of the industry. This practice has been applied across the industry and has been a driver in the development of more sensitive metabolism assays. The performance of the regression method was in a similar range in the external evaluation set when the compounds having significant extrahepatic metabolism (pimasertib and MS10) and enterohepatic recirculation (asimadoline) were removed, which supports its validity (Table 4). Recently, a number of methods allowing sampling over days have been introduced. It is plausible that the application of such methods may improve accuracy not only in the low clearance range (e.g. *in vitro* CL_{int} <1 μL/min/10^6 cells) but also better accuracy over the entire clearance range as longer incubation time allows for a better definition of CL_{int} values also at higher values (Bonn et al., 2016; Chan et al., 2013; Schaefer et al., 2016).

The 90% prospective interval was indicative of prospective performance for the regression equation model. More than 90% (13 of 14) of the compounds with simple hepatic metabolism did fall inside the interval, suggesting that the interval truly was prospective. The findings regarding the outlier compounds comes across as expected and most likely predictable from preclinical data in two of the three cases. The outlier asimadoline has been reported to show enterohepatic recirculation, which caused an overestimation of clearance. The mechanism, which was similar in animals and

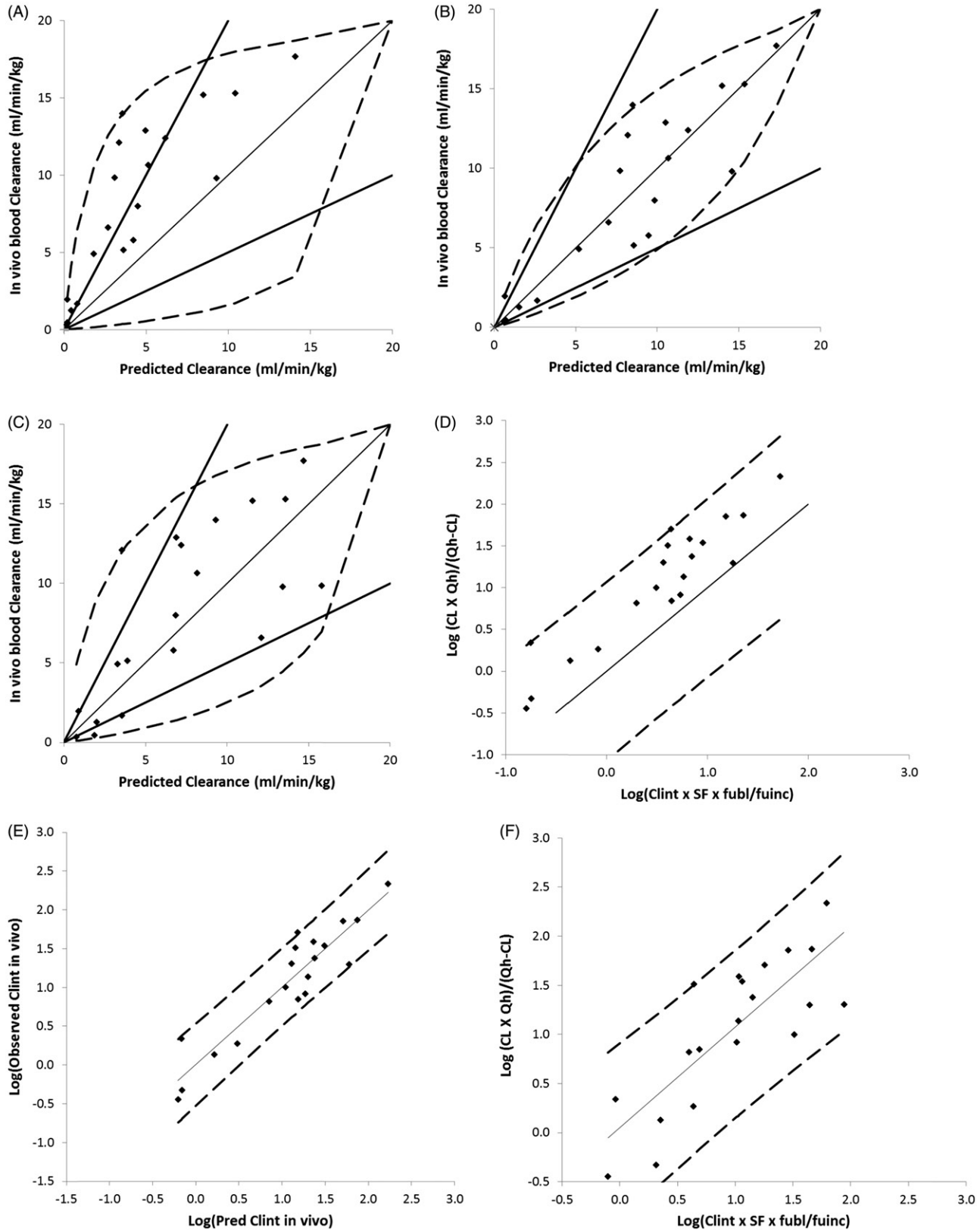

Figure 2. Comparison between predicted CL from hepatocytes for the reference compounds with known hepatic elimination (A) Direct scaling; (B) the regression equation method and (C) the Poulin et al. method. (D–F) Corresponding graphs are given in logarithms of measured and predicted CL_{int} *in vivo* for the direct scaling, the regression equation and the Poulin et al. methods, respectively. Dashed line, 90% confidence interval; solid line, twofold prediction error; thin line, line of identity.

human, involved intestinal hydrolysis of a glucuronide which was the major metabolite in animal bile (Camilleri, 2008). MS10 has a generally poor plasma stability and showed extrahepatic clearance in all the tested preclinical species. Pimasertib showed an excessive extrahepatic clearance exclusively in humans. This rather rare finding would make human risk assessment very difficult. The dataset does not contain any compounds that are biliary cleared.

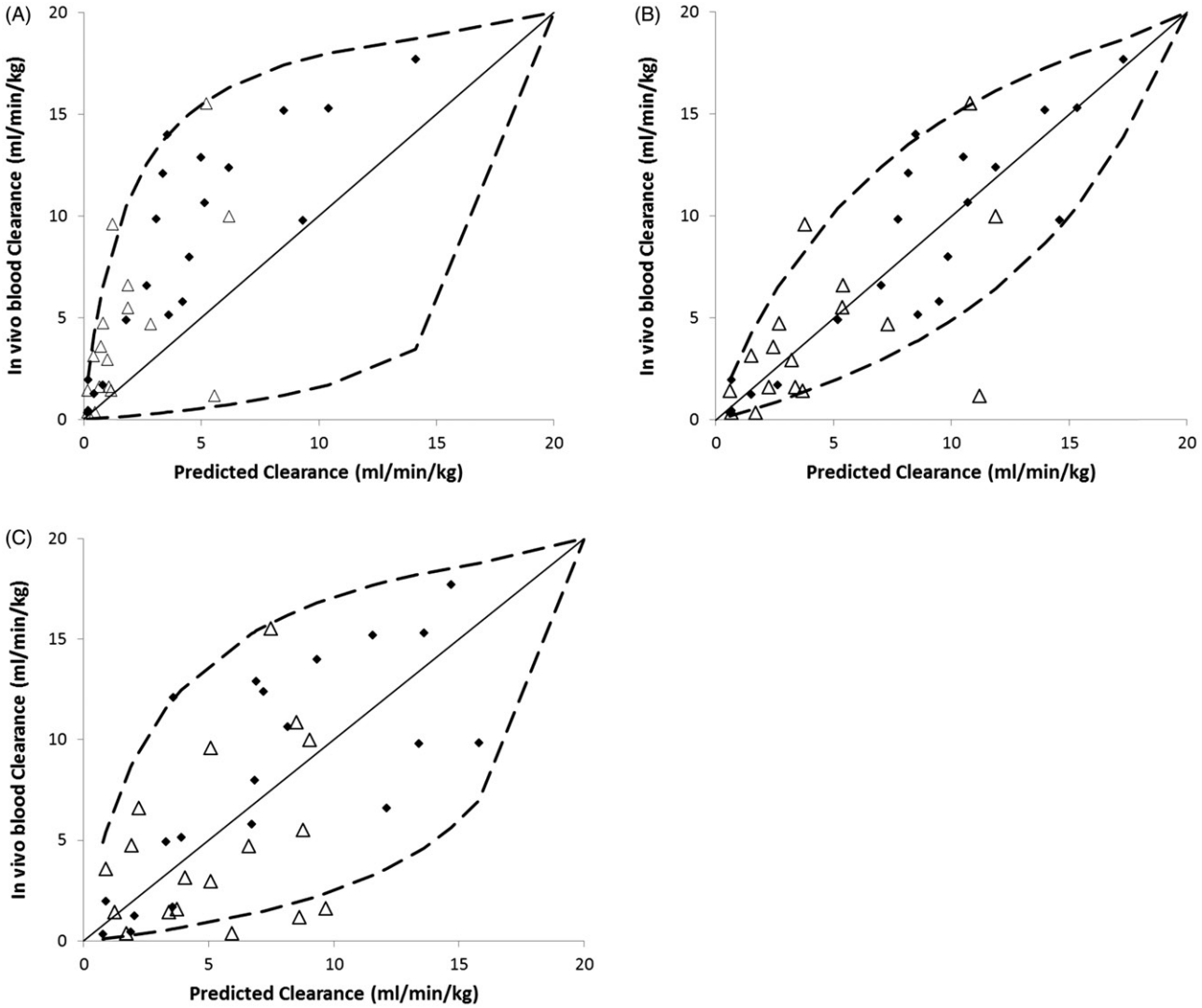

Figure 3. Comparison between predicted CL from hepatocytes for the reference compounds with known hepatic elimination and evaluation set compounds. Dashed line, 90% confidence interval; thin line, line of identity. Filled squares; reference compounds. Open triangles; evaluation set compounds. Panel (A), (B) and (C) refers to the direct scaling, regression equation method and the Poulin et al. method, respectively.

Risk assessment of this clearance pathways as well as renal clearance has been reported using an *in silico* model which may be used in addition to preclinical animal data (Hosey & Benet, 2015). Hence, these pathways do not necessarily pose a risk to human predictions. The findings of this study do still illustrate the danger of using these methods for compounds for which the major clearance processes are nonrelated to hepatic metabolism, i.e. extrahepatic metabolism and processes related to low permeability such as renal clearance. The 90% prospective interval, indicative of prospective performance for the ''Poulin et al.'' method was significantly wider than that for the regression equation model (Figure 3). As a consequence no compound with extrahepatic clearance fell outside of this range but the clearance of four compounds were underestimated to a degree that was not indicated by the reference compounds. Aside from asimadoline, for which the clearance process was complicated by enterohepatic recirculation, there were no obvious reason for the differences between the evaluation and reference sets.

This study indicated a better prediction accuracy for all endpoints when using the regression method versus the ''Poulin et al.'' method in the reference set as well as in the external evaluation set. The correlation coefficient was also better using regression than the ''Poulin et al.'' method (Table 4). This contrast to the findings by Poulin et al. (2012b). The difference may be explained by the composition of the datasets. An alternative explanation may be that Poulin used data from various sources but did choose to apply a correction from one of the labs providing data and thereby introducing bias causing a somewhat lower performance of the regression method. Given the fact that the regression method requires a calibration using reference compounds whereas the ''Poulin et al.'' method has a theoretical basis still makes the ''Poulin et al.'' method attractive, especially earlier in the value chain when wide prediction intervals are less of an issue. The potential complication of this approach, in that the binding partner has to be assigned which is a relatively resource intense endeavor, has been shown to be successfully mitigated by using a relatively simple assignment based on ion class (Poulin et al., 2012b).

The reported values in this study indicating an accuracy of 64% within twofold, a correlation coefficient of 0.75 and an

AAFE of 1.7 for the regression method in the external test set with many low clearance compounds (7 of 17 compounds had CL <10% of liver blood flow) suggests that this method is highly relevant. These findings contrast to a number of other studies that indicate the accuracy of *in vivo–in vitro* correlation was comparable to methods like allometry (Hosea et al., 2009). However, as datasets were not identical such comparisons are less straightforward. One may speculate around two possible reasons for the discrepancy: many studies neither correct for the well-known under estimation using IVIVE nor do they account for additional clearance pathways. As an example, although the PhRMA CPCDC Initiative discussed underestimation using IVIVE from microsomes as well as from hepatocytes, no method accounting for this phenomena was included in the analysis (Ring et al., 2011).

Based on this study, we conclude that the combined use of a well-calibrated consistent hepatocyte *in vitro* metabolism assay, the regression method and the associated prediction interval provides best practice for final estimate of human clearance to be used in dose prediction. This method has been implemented consistently into the drug discovery and preclinical development of Merck for all projects.

Acknowledgements

Tobias Feige is gratefully acknowledged for help with the clinical data analysis. Christian Stelz and Christiane Müller are gratefully acknowledged for technical assistance.

Declaration of interest

This work was supported by Merck. All authors were employed by Merck at the time this study was performed.

References

Armitage P, Berry G. (1994). Statistical methods in medical research. 3rd ed. Oxford: Blackwell Scientific Publications.

Bonn B, Svanberg P, Janefeldt A, et al. (2016). Determination of human hepatocyte intrinsic clearance for slowly metabolized compounds: comparison of a primary hepatocyte/stromal cell co-culture with plated primary hepatocytes and HepaRG. Drug Metab Dispos 44: 527–33.

Brandon EF, Raap CD, Meijerman I, et al. (2003). An update on *in vitro* test methods in human hepatic drug biotransformation research: pros and cons. Toxicol Appl Pharmacol 189:233–46.

Brown HS, Griffin M, Houston JB. (2007). Evaluation of cryopreserved human hepatocytes as an alternative in vitro system to microsomes for the prediction of metabolic clearance. Drug Metab Dispos 35: 293–301.

Camilleri M. (2008). Novel pharmacology: asimadoline, a kappa-opioid agonist, and visceral sensation. Neurogastroenterol Motil 20:971–9.

Chan TS, Yu H, Moore A, et al. (2013). Meeting the challenge of predicting hepatic clearance of compounds slowly metabolized by cytochrome P450 using a novel hepatocyte model, HepatoPac. Drug Metab Dispos 41:2024–32.

Cook D, Brown D, Alexander R, et al. (2014). Lessons learned from the fate of AstraZeneca's drug pipeline: a five-dimensional framework. Nat Rev Drug Discov 13:419–31.

Dolgos H, Trusheim M, Gross D, et al. (2016). Translational medicine guide transforms drug development processes: the recent Merck experience. Drug Discov Today 21:517–26.

Donato MT, Castell JV. (2003). Strategies and molecular probes to investigate the role of cytochrome P450 in drug metabolism: focus on in vitro studies. Clin Pharmacokinet 42:153–78.

Grime K, Riley RJ. (2006). The impact of in vitro binding on *in vitro-in vivo* extrapolations, projections of metabolic clearance and clinical drug-drug interactions. Curr Drug Metab 7:251–64.

Hallifax D, Galetin A, Houston JB. (2008). Prediction of metabolic clearance using fresh human hepatocytes: comparison with cryopreserved hepatocytes and hepatic microsomes for five benzodiazepines. Xenobiotica 38:353–67.

Hansch C, Leo A, Hoekman D. (1995). In exploring QSAR. hydrophobic, electronic and steric constants. Washington (DC): American Chemical Society.

Hewitt NJ, Lechón MJ, Houston JB, et al. (2007). Primary hepatocytes: current understanding of the regulation of metabolic enzymes and transporter proteins, and pharmaceutical practice for the use of hepatocytes in metabolism, enzyme induction, transporter, clearance, and hepatotoxicity studies. Drug Metab Rev 39:159–234.

Hosea NA, Collard WT, Cole S, et al. (2009). Prediction of human pharmacokinetics from preclinical information: comparative accuracy of quantitative prediction approaches. J Clin Pharmacol 49: 513–33.

Hosey CM, Benet LZ. (2015). Predicting the extent of metabolism using *in vitro* permeability rate measurements and in silico permeability rate predictions. Mol Pharm 12:1456–66.

Houston JB. (1994). Utility of in vitro drug metabolism data in predicting in vivo metabolic clearance. Biochem Pharmacol 47: 1469–79.

Ito K, Houston JB. (2004). Comparison of the use of liver models for predicting drug clearance using in vitro kinetic data from hepatic microsomes and isolated hepatocytes. Pharm Res 21:785–92.

Ito K, Houston JB. (2005). Prediction of human drug clearance from in vitro and preclinical data using physiologically based and empirical approaches. Pharm Res 22:103–12.

Kilford PJ, Gertz M, Houston JB, Galetin A. (2008). Hepatocellular binding of drugs: correction for unbound fraction in hepatocyte incubations using microsomal binding or drug lipophilicity data. Drug Metab Dispos 36:1194–7.

Kovar A, Peters T, Beier N, Derendorf H. (2001). Pharmacokinetic/pharmacodynamic evaluation of the NHE inhibitor eniporide. J Clin Pharmacol 41:139–48.

Krösser S, Wolna P, Fischer TZ, et al. (2015). Effect of ceralifimod (ONO-4641) on lymphocytes and cardiac function: randomized, double-blind, placebo-controlled trial with an open-label fingolimod arm. J Clin Pharmacol 55:1051–60.

Luzier A, Morse GD. (1993). Intravascular distribution of zidovudine: role of plasma proteins and whole blood components. Antiviral Res 21:267–80.

McGinnity DF, Soars MG, Urbanowicz RA, Riley RJ. (2004). Evaluation of fresh and cryopreserved hepatocytes as in vitro drug metabolism tools for the prediction of metabolic clearance. Drug Metab Dispos 32:1247–53.

Nordell P, Svanberg P, Bird J, Grime K. (2013). Predicting metabolic clearance for drugs that are actively transported into hepatocytes: incubational binding as a consequence of in vitro hepatocyte concentration is a key factor. Drug Metab Dispos 41:836–43.

Obach RS. (1999). Prediction of human clearance of twenty-nine drugs from hepatic microsomal intrinsic clearance data: an examination of in vitro half-life approach and nonspecific binding to microsomes. Drug Metab Dispos 27:1350–9.

Poulin P, Haddad S. (2013). Toward a new paradigm for the efficient in vitro-in vivo extrapolation of metabolic clearance in humans from hepatocyte data. J Pharm Sci 102:3239–51.

Poulin P, Cornelis Hop ECA, Ho Q, et al. (2012a). Comparative assessment of in vitro–in vivo extrapolation methods used for predicting hepatic metabolic clearance of drugs. J Pharm Sci 101:4308–26.

Poulin P, Kenny JR, Eca H, Haddad S. (2012b). In vitro–in vivo extrapolation of clearance: modeling hepatic metabolic clearance of highly bound drugs and comparative assessment with existing methods. J Pharm Sci 101:838–51.

Ravandi F, Pigneux A, DeAngelo DJ, et al. (2015). Clinical, pharmacokinetic and pharmacodynamic data for the MEK1/2 inhibitor pimasertib in patients with advanced hematologic malignancies. Blood Cancer J 5:e375.

Richard B, Launay-Iliadis MC, Iliadis A, et al. (1992). Pharmacokinetics of mitoxantrone in cancer patients treated by high-dose chemotherapy and autologous bone marrow transplantation. Br J Cancer 65: 399–404.

Riley RJ, McGinnity DF, Austin RP. (2005). A unified model for predicting human hepatic, metabolic clearance from in vitro intrinsic clearance data in hepatocytes and microsomes. Drug Metab Dispos 33:1304–11.

Ring BJ, Chien JY, Adkison KK, et al. (2011). PhRMA CPCDC initiative on predictive models of human pharmacokinetics, part 3: comparative assessement of prediction methods of human clearance. J Pharm Sci 100:4090–110.

Schaefer M, Schänzle G, Bischoff D, Süssmuth RD. (2016). Upcyte human hepatocytes: a potent in vitro tool for the prediction of hepatic clearance of metabolically stable compounds. Drug Metab Dispos 44: 435–44.

Scheible H, Kraetzer F, Johne A, Wimmer E. (in manuscript). Metabolism of the MEK1/2 inhibitor pimasertib involves a novel conjugation with phospho ethanolamine in patients with solid tumors.

Soars MG, Grime K, Sproston JL, et al. (2007). Use of hepatocytes to assess the contribution of hepatic uptake to clearance in vivo. Drug Metab Dispos 35:859–65.

Sohlenius-Sternbeck AK, Afzelius L, Prusis P, et al. (2010). Evaluation of the human prediction of clearance from hepatocyte and microsome intrinsic clearance for 52 drug compounds. Xenobiotica 40:637–49.

Sohlenius-Sternbeck AK, Jones C, Ferguson D, et al. (2012). Practical use of the regression offset approach for the prediction of in vivo intrinsic clearance from hepatocytes. Xenobiotica 42: 841–53.

Stringer R, Nicklin PL, Houston JB. (2008). Reliability of human cryopreserved hepatocytes and liver microsomes as in vitro systems to predict metabolic clearance. Xenobiotica 38:1313–29.

Tang H, Hussain A, Leal M, et al. (2007). Interspecies prediction of human drug clearance based on scaling data from one or two animal species. Drug Metab Dispos 35:1886–93.

Williams JA, Hyland R, Jones BC, et al. (2004). Drug-drug interactions for UDP-glucuronosyltransferase substrates: a pharmacokinetic explanation for typically observed low exposure (AUCi/AUC) ratios. Drug Metabol Dispos 32:1201–8.