Original scientific paper

In silico ADME in drug design – enhancing the impact

Susanne Winiwarter[1],*, Ernst Ahlberg[1], Edmund Watson[1], Ioana Oprisiu[1], Mickael Mogemark[1], Tobias Noeske[1], Nigel Greene[2]

[1]*Drug Safety and Metabolism, IMED Biotech Unit, AstraZeneca, Gothenburg, Sweden*
[2]*Drug Safety and Metabolism, IMED Biotech Unit, AstraZeneca, Boston, US*

*Corresponding Author: E-mail: susanne.winiwarter@astrazeneca.net

Abstract

Each year the pharmaceutical industry makes thousands of compounds, many of which do not meet the desired efficacy or pharmacokinetic properties, describing the absorption, distribution, metabolism and excretion (ADME) behavior. Parameters such as lipophilicity, solubility and metabolic stability can be measured in high throughput in vitro assays. However, a compound needs to be synthesized in order to be tested. In silico models for these endpoints exist, although with varying quality. Such models can be used before synthesis and, together with a potency estimation, influence the decision to make a compound. In practice, it appears that often only one or two predicted properties are considered prior to synthesis, usually including a prediction of lipophilicity. While it is important to use all information when deciding which compound to make, it is somewhat challenging to combine multiple predictions unambiguously. This work investigates the possibility of combining in silico ADME predictions to define the minimum required potency for a specified human dose with sufficient confidence. Using a set of drug discovery compounds, in silico predictions were utilized to compare the relative ranking based on minimum potency calculation with the outcomes from the selection of lead compounds. The approach was also tested on a set of marketed drugs and the influence of the input parameters investigated.

Keywords

Drug discovery; pharmacokinetics; *in silico* predictions; multi-parameter ranking

Introduction

Drug design is a multi-parameter optimization process, with pharmacokinetic properties describing absorption, distribution, metabolism and excretion (ADME) of a compound being important to consider early on. These parameters describe the pharmacokinetics of a compound and are key to determining the dose required for efficacy. Parameters such as lipophilicity, solubility, and metabolic stability can be measured in a high throughput manner *in vitro* and are, thus, often used as early ADME screens. During lead identification and lead optimization phases a substantial number of compounds is synthesized and characterized in such screening assays for selection towards additional, more costly *in vitro* and *in vivo* experiments to be able to finally select a single candidate drug. However, a compound needs to be physically available to be subjected to such assays. Thus, drug industry generates thousands of compounds per month out of which many do not show desirable ADME properties.

In silico models for such ADME endpoints, on the other hand, have been available for a long time, though with varying quality and usability [1-6]. Ideally, such models would be used before synthesis and, together with any potency estimation for the specific case, influence the decision to make a compound. Consequent usage of such *in silico* predictions has the potential to considerably reduce the number of synthesized compounds with inadequate ADME properties. In practice, it seems that most often only a few predicted properties are really considered before synthesis, usually including lipophilicity as important parameter. While it is understood that all available knowledge should be used to select which compound to make, it is not easy to define how the outcome of various predictions can be combined unambiguously. Lately, the use of multi-parameter optimization and scoring tools has been proposed and shown to be of value [7, 8]. However, the definition of the scoring functions can be difficult and may be a bit arbitrary. As physiologically meaningful scoring function the predicted dose to man (D2M) has been suggested to utilize as early as possible. Several reports validated early D2M predictions both from preclinical *in vivo* data [9] as well as from *in vitro* data combined with *in silico* predictions [10]. However, using *in silico* predictions only was not found to be reliable enough, since both potency and pharmacokinetic properties would need to be accurately predicted.

Here we concentrate on the pharmacokinetic properties only and suggest to use *in silico* ADME models to predict what potency would be required for a specific compound to enable coverage during the whole dosing interval. By reverting the D2M equation and setting the target dose to, for example, 100 mg once daily, the achievable plasma concentration can be calculated from *in silico* predicted parameters and the minimum required potency ("threshold pIC50") defined as ranking score for virtual compounds. We show how such a ranking score could be used in the discovery setting on the example of a set of 27 compounds leading towards a candidate structure [11]. We also test the concept on a set of known drugs with information about dosing and human pharmacokinetics [9] and compare the results from purely *in silico* predictions to those from *in vitro* data and from human *in vivo* data. We also investigate how the *in silico* derived parameters influence the outcome of the equations, highlighting concepts to help drug discovery to improve pharmacokinetic properties during design.

Experimental

Inverse dose-to-man (D2M) prediction

Equations used are standard pharmacokinetic equations [9, 12, 13] based on a one compartment model considering immediate absorption to estimate the average (C_{ave}, Eq. 1), minimum (C_{min}, Eq. 2) and maximum (C_{max}, Eq. 3) concentration of a drug at steady state.

$$C_{ave} = \frac{F \cdot D_{po} \cdot {^\tau}/_{24}}{Cl \cdot \tau} \tag{1}$$

$$C_{min} = \frac{F \cdot D_{po} \cdot {^\tau}/_{24}}{V \cdot (e^{k_e \cdot \tau} - 1)} \tag{2}$$

$$C_{max} = \frac{F \cdot D_{po} \cdot {^\tau}/_{24}}{V \cdot (1 - e^{-k_e \cdot \tau})} \tag{3}$$

where F is bioavailability (Eq. 4), D_{po} is the daily oral dose, Cl is the plasma clearance estimated using the well-stirred model (Eq. 5), τ is the dosing interval, typically set to 24 h, V is the volume of distribution (predicted) and k_e is the elimination constant (Eq. 6).

$$F = F_{abs} \cdot F_g \cdot F_h \, , \tag{4}$$

where F_{abs} (fraction absorbed) is estimated from permeability (Eq. 7), F_g (fraction escaping gut metabolism) is set to 1 and F_h (fraction escaping hepatic metabolism) is estimated from Cl (Eq. 8).

$$Cl_b = \frac{Q_h \cdot CLint_{in\ vivo} \cdot fu_b}{Q_h + CLint_{in\ vivo} \cdot fu_b} \qquad (5)$$

with Cl_b being blood clearance, Q_h the liver blood flow (20 ml/min/kg), $CLint_{in\ vivo}$ the *in vivo* intrinsic clearance estimated from *in vitro CLint* (Eq. 9), and fu_b the fraction unbound in blood.

$$Cl = Cl_b \cdot R_b , \qquad (5a)$$

where R_b is the blood plasma ratio and set to 1 for neutrals and bases, whereas set to 0.55 for acids and zwitterions.

$$fu_b = fu/R_b , \qquad (5b)$$

where fu is fraction unbound in plasma (predicted).

$$k_e = Cl/V \qquad (6)$$

$$F_{abs} = 1,\ \text{if } P_{app} > 1 \times 10^{-6}\ \text{cm/s, otherwise } F_{abs} = P_{app} , \qquad (7)$$

where P_{app} is apparent permeability measured in Caco2 cells, utilising the intrinsic Caco2 permeability assay [14] (predicted).

$$F_h = 1 - Cl_b/Q_h \qquad (8)$$

$$CLint_{in\ vivo} = CF \cdot CLint_{sc}, \qquad (9)$$

where CF is an empirical correction factor adopted for the simplified regression offset approach (CF=3, see below), and $CLint_{sc}$ is the scaled intrinsic clearance from the human hepatocyte incubation (Eq. 9a).

$$CLint_{sc} = CLint_{HH} \cdot SF_{HH}/fu_{inc}, \qquad (9a)$$

where $CLint_{HH}$ is the intrinsic clearance in human hepatocytes (predicted), SF_{HH} are the scaling factors (120 million cells per g liver x 1680 g liver weight / 70 kg body weight = 2.88) and fu_{inc} is the fraction unbound in the hepatocyte incubation (predicted).

The minimum potency value, e.g., "threshold pIC50" based on C_{min}, is then calculated as the negative logarithm of the predicted concentration considering the specific fold coverage, typically set to 1 or 3 (Eq. 10).

$$\text{threshold pIC50} = -\log(C_{min}/\text{fold coverage}) \qquad (10)$$

Simplified regression offset approach – correction factor (CF)

The regression offset approach [15] was shown to improve *in vitro* to *in vivo* scaling of clearance, especially when experiments where compared from different laboratories (sites) and using different batches of hepatocytes or microsomes. Extensive in-house data analysis (data not shown) indicated that overall the regression slope can be set to 1 and the offset to about 0.5, giving an equivalent level of correct predictions in three species, human, rat and dog. Thus, *in vivo CLint* can be estimated by multiplying the scaled *CLint* with a correction factor of 3. Here we adopt this procedure to estimate *in vivo CLint* directly from *in silico* predictions.

In silico models

Five *in silico* models were used as input for the above calculations. Human volume of distribution at

steady state, intrinsic Caco2 permeability, human plasma protein binding, human hepatocyte intrinsic clearance, and fraction unbound in the hepatocyte incubation. All but the volume of distribution model used AstraZeneca in-house experimental data. All models are available within AstraZeneca.

1. Human volume of distribution (*V)*

Human volume of distribution data for about 700 compounds were collated from literature [16-18] and randomly split into a training and test set (n = 544 and 144, respectively). Both acids, bases, neutral and zwitterionic compounds are included in both datasets. It is a random forest model [19] using physico-chemical descriptors [20] including ACD log P and log D [21] and clog P [22]. The experimental data is modelled in its logarithmic form. The prediction explains ~70 % of the variance of the test set, and has an error in prediction of 0.4 log units.

2. Intrinsic Caco2 permeability (P_{app})

The model is built on in-house data of intrinsic Caco2 permeability, apparent permeability measured in Caco2 cells in presence of a defined transport inhibitor cocktail as described earlier [14], and updated about every sixth month. The present training set consists of >4,000 data points. About 100 simple physico-chemical descriptors [20] including ACD log P and log D [21] and clog P [22] are calculated from the compound structures. The data is modelled as log P_{app} using random forest regression [19] as implemented in scikit-learn [23]. The latest temporal test set comprises about 300 compounds and shows an R^2 of 0.4 and a root mean squared error of predictions (RMSEP) of 0.6 (log scale), with about 60 % of the compounds being within 3-fold of the experimental value. The model can be used to distinguish between high and low permeability compounds with classification accuracy above 0.8.

3. Human plasma protein binding as fraction unbound (*f*u*)*

The model is built on in-house data of human plasma protein binding generated using equilibrium dialysis in high throughput assays as described earlier [24-26]. It is updated monthly. The present training set consists of almost 90,000 compounds. The data is modelled as log K (= log[(fraction unbound)/(fraction bound)]). The modelling procedure utilizes support vector machines [27] with a linear kernel, signature descriptors [28] and the conformal prediction framework [29] as implemented in cpsign from GenettaSoft [30]. The latest temporal test set comprising 750 compounds shows an R^2 of 0.7 and an RMSEP of 0.4, with about 80 % of the compounds being within 3-fold of the experimental value of *f*u.

4. Human hepatocyte intrinsic clearance (*CLint*$_{HH}$)

The model is built on in-house data of human hepatocyte intrinsic clearance generated in high throughput assays using incubations of either cryopreserved or fresh human hepatocytes at 37 °C for up to 120 min as described earlier [26, 31, 32]. It is updated about every sixth month. The present training set contains more than 11,000 compounds and the data is modelled as log *CLint*. The modelling procedure utilizes support vector machines [27] with a linear kernel, signature descriptors [28] and the conformal prediction framework [29] as implemented in cpsign from GenettaSoft [30]. The latest temporal test set comprising almost 200 compounds shows an R^2 of 0.2 and an RMSEP of 0.4, with ~75 % of the compounds being within 3-fold of the experimental *CLint* value.

5. Fraction unbound in the hepatocyte incubation (*f*u$_{inc}$)

The model is built on in-house binding data measured in cryopreserved rat hepatocyte incubations as described earlier [33]. The model is updated approximately every sixth month and the present data set contains about 1,700 compounds. The data is modelled as log K (= log[(fraction unbound)/(fraction

bound)]). The modelling procedure utilizes support vector machines [27] with a radial basis function kernel, signature descriptors [28] and the conformal prediction framework [29] as implemented in cpsign from GenettaSoft [30]. The latest temporal test set comprising about 200 compounds shows an R^2 of 0.5 and an RMSEP of 0.5, with ~75 % of the compounds being within 3-fold of the experimental fu_{inc} value.

Global Sensitivity Analysis

Global sensitivity analysis was performed by a quasi-Monte Carlo method using the Fourier amplitude sensitivity test (FAST) and Sobol' sensitivity, which are implemented in the Global sensitivity analysis toolbox (GSAT) [34] in MATLAB [35] using 20,000 sample points.

Local Sensitivity Analysis

Local sensitivity analysis was performed by taking Jacobian matrices of the model with respect to the model parameters and other parameter points [36]. These matrices contain actual parameter values which can be examined, or plotted graphically against other parameters, to show which have the greatest effect at each point.

$$S(q,p) = \begin{pmatrix} \dfrac{\partial x(q_1,p)}{\partial p_1} & \cdots & \dfrac{\partial x(q,p)}{\partial p_n} \\ \vdots & \ddots & \vdots \\ \dfrac{\partial x(q_m,p)}{\partial p_1} & \cdots & \dfrac{\partial x(q_m,p)}{\partial p_n} \end{pmatrix}, \tag{11}$$

where $q = (q_1, \ldots, q_m)$ is a vector of parameter points to be examined, $p = (p_1, \ldots, p_n)$ is a vector of other parameters which are being examined, and x is a scalar output from the model. This was performed in MATLAB [35] using automatic differentiation of the parameters via the myAD toolkit [37].

Data sets

Two data sets were used for the present analysis. Data set 1 was a set of 27 drug discovery compounds exemplifying important compounds leading towards candidate selection [11]. Data set 2 was a set of 21 marketed drugs, for which potency data, dosing information and human pharmacokinetic data had been collated and used to show feasibility of early D2M predictions [9].

Results and Discussion

Calculation of required minimum potency (threshold pIC50) for a set of drug discovery compounds

As proof of concept a set of drug discovery molecules leading towards a candidate structure [11] was evaluated. Intrinsic clearance in human hepatocytes, fraction unbound in the incubation, fraction unbound in plasma, volume of distribution and intrinsic Caco2 permeability were predicted for all compounds using the present in-house models (see Table 1) and minimum and maximum plasma concentrations for a once daily oral dose of 100 mg calculated using equations 2 and 3. The project aimed at minimum plasma concentrations with 3-fold coverage over potency measured as pIC50 in a whole blood assay. Thus, threshold pIC50 values were calculated from $C_{min,total}/3$ (see Table 2). Blood plasma ratios for these mainly neutral compounds were considered to be 1. Protein binding was considered in the scaling approach, but threshold pIC50s were based on resulting total plasma concentration to match the potency measurement.

The data indicated that some of the compounds had very short half-life and thus a once daily dose resulted in a high threshold pIC50 to cover potency over 24 hours and a high C_{max}/C_{min} ratio. Using a 10 times higher dose would reduce the threshold pIC50 by one unit, but show the same C_{max}/C_{min} ratio, since the applied model assumes linear kinetics. Considering only compounds with a threshold pIC50 below 9 and a C_{max}/C_{min} ratio below 100 about 12 compounds remained. This set of compounds included compound **15b**, the compound finally selected, and three of the remaining four compounds of higher interest for which rat *in vivo* results were reported. Compound **22**, which according to the present analysis was least favourable, was actually found in the rat study to have short half-life and, thus, not suitable to progress.

In summary, it seems that for this data set the threshold pIC50 together with the predicted C_{max}/C_{min} ratio could be used for ranking the compounds and correctly identify those with inferior pharmacokinetics. The results also highlight the danger in deprioritizing compounds just by applying a cut-off for, for example, metabolic stability. Here, three compounds had (predicted) *CLint* values higher than 20 µl/min/10^6 cells, compounds **12**, **14a** and **19**. For **12**, a reference compound, threshold pIC50 was one of the lowest in the whole set based on a predicted blood clearance below 5 ml/min/kg, whereas **14a** showed an intermediate threshold pIC50 (8.2) and blood clearance just above 5 ml/min/kg. Only for compound **19** the threshold pIC50 was estimated to be on the high end within the present set, even though the blood clearance was again predicted as just above 5 ml/min/kg. **22**, the compound with the highest threshold pIC50 has almost double the blood clearance despite a lower hepatocyte *CLint* (~ 8 µl/min/10^6 cells). Rank order differences between *in vitro* (or *in silico*) *CLint* and predicted *in vivo* clearance can be explained by the different binding properties: High binding in the incubation, i.e., low fu_{inc} values, will potentiate the metabolic instability measured in the incubation, whereas high binding in plasma, i.e., low fu, can to some extent mitigate low metabolic stability seen *in vitro*. Furthermore, minimum concentrations mainly depend on a compound's half-life, determined by clearance and volume of distribution. Thus, volume of distribution is another important factor that again may change the rank order.

Table 1. Predicted properties for compound set 1 [11] using AstraZeneca in-house *in silico* ADME models

cmpd ID[a]	$CLint_{HH}$[b] (μl/min/10^6 cells)	fu_{inc}[c] %free	fu[d] %free	P_{app}[e] (10^{-6} cm/sec)	V[f] (L/kg)	MW[g] (g/mol)	ion class
12	20.2	37.7	1.1	20	2.0	447.6	neutral
13a	15.0	38.1	2.5	5	2.1	433.6	neutral
13b	6.5	25.4	3.4	9	2.0	433.6	neutral
14a	27.3	80.2	2.4	26	1.7	447.5	neutral
14b	7.5	72.4	15.3	5	2.0	446.5	base
14c	9.5	73.1	10.5	36	1.3	431.5	neutral
14d	10.3	83.5	9.4	10	0.9	444.5	neutral
14e	6.8	74.9	4.3	27	1.5	447.5	neutral
15a	10.3	70.6	7.5	2	1.2	433.5	neutral
15b	3.7	56.9	10.2	7	1.4	433.5	neutral
15c	6.7	55.7	5.3	12	1.5	432.5	neutral
15d	3.6	47.4	10.7	7	1.4	433.5	neutral
15e	6.3	55.9	4.2	11	1.4	447.5	neutral
15f	6.1	47.6	5.0	7	1.4	431.5	neutral
19	32.3	70.9	1.8	10	0.8	416.4	neutral
20	8.1	76.1	4.3	12	1.0	458.5	neutral
21	12.9	76.0	2.5	13	1.2	415.5	neutral
22	7.9	39.2	10.3	6	0.9	386.4	neutral
29	4.7	49.8	8.0	9	1.5	447.5	neutral
30	5.8	47.5	6.7	9	1.6	447.5	neutral
31	4.1	47.0	9.6	10	1.5	447.5	neutral
32	5.4	52.2	11.5	9	1.5	445.5	neutral
37	2.6	40.6	8.8	8	1.3	469.5	neutral
41	3.7	56.9	10.2	7	1.4	433.5	neutral
42	4.6	46.6	7.7	8	1.7	461.6	neutral
47	2.3	67.1	24.2	7	1.3	433.5	neutral
49	4.0	49.0	16.9	8	1.4	421.5	neutral

[a]cmpd ID: compound identification as originally defined [11]; [b]$CLint_{HH}$: intrinsic clearance in human hepatocytes; [c]fu_{inc}: fraction unbound in rat hepatocyte incubation; [d]fu: fraction unbound in human blood plasma; [e]P_{app}: apparent permeability in Caco2 cells with transport inhibition; [f]V: human volume of distribution at steady state; [g]MW: molecular weight

Table 2. Predicted PK properties for compound set 1 [11] based on *in silico* ADME model predictions, plasma concentrations calculated for a once daily dose of 100 mg (1.4 mg/kg), sorted by threshold pIC50

cmpd ID[a]	Cl blood[b] (ml/min/kg)	C_{ave}[c] (μM)	C_{min}[d] (μM)	C_{max}[e] (μM)	threshold pIC50 from C_{min}[f]	C_{max}/C_{min}[g]
14e	2.88	0.66	0.12904	1.90	7.4	14.7
12	4.17	0.42	0.06645	1.32	7.7	19.9
15e	3.41	0.54	0.06363	1.90	7.7	29.8
21	3.13	0.64	0.06105	2.44	7.7	39.9
15c	4.33	0.42	0.02902	1.73	8.0	59.5
20	3.30	0.55	0.02517	2.56	8.1	101.6
42	4.92	0.33	0.02259	1.38	8.1	61.0
13b	5.45	0.31	0.02254	1.25	8.1	55.4
37	3.96	0.43	0.02136	1.96	8.1	91.5
15b	4.47	0.40	0.01900	1.84	8.2	96.8
41	4.47	0.40	0.01900	1.84	8.2	96.8
15f	4.37	0.41	0.01864	1.93	8.2	103.4
14a	5.22	0.31	0.01757	1.39	8.2	79.2
13a	6.03	0.26	0.01675	1.13	8.3	67.7
29	4.88	0.34	0.01431	1.64	8.3	114.7
30	5.27	0.31	0.01154	1.53	8.4	132.3
31	5.36	0.30	0.00998	1.54	8.5	154.0
15d	5.26	0.32	0.00709	1.78	8.6	250.9
47	5.34	0.31	0.00599	1.80	8.7	300.1
14b	8.10	0.16	0.00259	0.97	9.1	374.3
32	6.83	0.21	0.00171	1.45	9.2	846.0
15a	6.41	0.24	0.00074	1.91	9.6	2573.8
49	7.51	0.20	0.00060	1.53	9.7	2537.8
14c	7.44	0.19	0.00046	1.58	9.8	3398.1
19	5.18	0.34	0.00045	3.01	9.8	6751.1
14d	6.65	0.22	0.00004	2.46	10.9	58276.0
22	9.46	0.14	0.00000	2.20	12.8	4832468.9

[a]cmpd ID: compound identification as defined in ref [11]; [b]Cl blood: predicted human blood clearance; [c]C_{ave}: predicted total average plasma concentration at steady state; [d]C_{min}: predicted total minimum plasma concentration at steady state; [e]C_{max}: predicted total maximum plasma concentration at steady state; [f]threshold pIC50 from C_{min}: minimum required pIC50 estimated from $C_{min}/3$ for a 3-fold coverage over 24 hours; [g]C_{max}/C_{min}: predicted ratio between maximum and minimum concentration at steady state

Calculation of threshold potency for a set of known drugs

As second test, we investigated the approach for a set of marketed drugs where information on dosing and human pharmacokinetic data had been collated by McGinnity et al. [9]. This dataset was previously used to evaluate the validity of early dose to man predictions [9, 10]. Table 3 shows *in silico* ADME parameters for the 21 drugs predicted with the present models at AstraZeneca.

Table 3. Predicted properties for 21 known drugs [9] using AstraZeneca in-house *in silico* ADME models

Drug name	$CLint_{HH}$[a] (µl/min/10^6 cells)	fu_{inc}[b] %free	fu[c] %free	P_{app}[d] (10^{-6} cm/sec)	V[e] (L/kg)	MW[f] (g/mol)	ion class
Acebutolol	7.7	79.0	80.8	2.5	1.5	336.4	base
Atenolol	1.6	89.7	92.4	0.3	1.5	266.3	base
Betaxolol	10.0	82.1	51.0	17.6	3.8	307.4	base
Bisoprolol	9.5	82.1	63.8	12.6	2.3	325.4	base
Carvedilol	25.0	19.5	0.6	4.8	1.9	406.5	base
Cimetidine	5.2	77.7	81.2	4.8	1.0	252.3	base
Cyclosporin	738.4	97.4	0.5	1.4	2.0	1203	neutral
Desloratidine	9.2	11.9	7.8	19.4	7.8	310.8	base
Diazepam	6.5	14.6	3.7	55.3	1.1	284.7	neutral
Diclofenac	55.2	52.8	1.8	33.8	0.3	296.2	acid
Famotidine	2.5	77.6	78.1	0.7	0.9	337.5	base
Indomethacin	7.5	39.7	0.6	74.1	0.2	357.8	acid
Irbesartan	15.1	17.7	0.8	15.6	0.9	428.5	acid
Metoprolol	4.9	87.3	73.1	13.4	2.4	267.4	base
Nadolol	4.4	79.8	76.9	0.4	2.1	309.4	base
Nifedipine	14.3	30.7	3.5	45.0	1.2	346.3	neutral
Pirenzepine	8.5	86.8	50.7	0.5	1.7	351.4	base
Prazosin	6.5	58.8	9.9	10.1	0.8	383.4	neutral
Propranolol	9.5	59.5	18.8	31.0	3.0	259.3	base
Ranitidine	5.6	69.4	80.2	2.2	1.2	314.4	base
Ritonavir	75.3	19.1	1.1	4.0	0.9	721	neutral

[a]$CLint_{HH}$: intrinsic clearance in human hepatocytes; [b]fu_{inc}: fraction unbound in rat hepatocyte incubation; [c]fu: fraction unbound in human blood plasma; [d]P_{app}: apparent permeability in Caco2 cells with transport inhibition; [e]V: human volume of distribution at steady state; [f]MW: molecular weight.

To take into account that these drugs are not necessarily given once daily only, we adopted the dosing regimen suggestion by McGinnity et al. [9] when calculating the minimum required potency, threshold pX (see Table 4). The resulting threshold pX values are below 10 for all but four of the compounds, thus indicating that the approach is able to estimate pharmacokinetic behaviour to some extent. Additionally, the C_{max}/C_{min} ratio was estimated as below 100 for most of the compounds, even though it showed values above 1000 for three of them (ritonavir, bisoprolol and diclofenac). The availability of both *in vitro* [10] and *in vivo* [9] data for the compounds prompted us to investigate, to what extent the usage of experimental data would change the results (see Table 5 and Figure 1).

Table 4. Dosing interval, predicted plasma concentrations and threshold pX based on a daily dose of 100 mg (1.4 mg/kg), given as 1-6 doses per day, for 21 known drugs [9]

Drug name	Dosing interval[a] (h)	$C_{min,total}$[b] (μM)	$C_{min,free}$[c] (μM)	$C_{max,total}$[d] (μM)	threshold pX from $C_{min,free}$[e]	C_{max}/C_{min}[f]
Acebutolol	8	0.00165	0.00134	0.21153	8.9	127.9
Atenolol	12	0.00742	0.00686	0.35626	8.2	48.0
Betaxolol	12	0.01102	0.00563	0.17910	8.2	16.3
Bisoprolol	24	0.00004	0.00002	0.44702	10.6	11795.4
Carvedilol	8	0.19649	0.00114	0.65630	8.9	3.3
Cimetidine	4	0.01081	0.00878	0.28422	8.1	26.3
Cyclosporin	12	0.00155	0.00001	0.12384	11.2	79.9
Desloratidine	24	0.01218	0.00095	0.17440	9.0	14.3
Diazepam	12	0.00572	0.00021	1.35487	9.7	236.9
Diclofenac	4	0.00002	0.00000	1.26227	12.4	55501.6
Famotidine	6	0.00631	0.00493	0.42939	8.3	68.1
Indomethacin	6	0.16284	0.00095	5.90360	9.0	36.3
Irbesartan	8	0.01792	0.00015	0.82347	9.8	46.0
Metoprolol	8	0.02235	0.01635	0.29316	7.8	13.1
Nadolol	8	0.00547	0.00421	0.10158	8.4	18.6
Nifedipine	6	0.04706	0.00164	0.54780	8.8	11.6
Pirenzepine	12	0.00055	0.00028	0.18872	9.6	345.6
Prazosin	6	0.04517	0.00448	0.84533	8.3	18.7
Propranolol	8	0.05111	0.00961	0.32040	8.0	6.3
Ranitidine	6	0.00271	0.00218	0.26131	8.7	96.3
Ritonavir	8	0.00021	0.00000	0.26073	11.6	1239.4

[a]Dosing interval as suggested by McGinnity et al. [9], based on their predicted $t_{1/2}$: $t_{1/2}$>8h → 24h; $t_{1/2}$ 4-8 h → 12h; $t_{1/2}$ 2-4h → 8h; $t_{1/2}$ 1-2 h → 6h; $t_{1/2}$ < 1h → 4h; [b]$C_{min,total}$: predicted minimum total concentration at steady state; [c]$C_{min,free}$: predicted minimum concentration at steady state corrected for plasma protein binding; [d]$C_{max,total}$: predicted maximum total concentration at steady state; [e]threshold pX from $C_{min,free}$: required minimum pX estimated from free C_{min} for coverage over dosing interval; [f]C_{max}/C_{min}: predicted ratio between maximum and minimum total concentration at steady state.

Table 5. Threshold pX based on a daily dose of 100 mg (1.4 mg/kg), given as 1-6 doses per day, for 21 known drugs[9], derived from *in silico* data (as above), *in vitro* data [9, 10, 14], and actual human *in vivo* data [9]

Drug name	Threshold pX from *in silico*[a]	Threshold pX from *in vitro*[b]	Threshold pX from *in vivo*[c]
Acebutolol	8.9	6.9[d]	7.8
Atenolol	8.2	7.8[e,f]	6.5
Betaxolol	8.2	7.6[e]	6.4
Bisoprolol	10.6	7.6[e]	6.6
Carvedilol	8.9	9.7[d]	9.3
Cimetidine	8.1	6.8[d,f]	7.1
Cyclosporin	11.2	NV[g]	8.9
Desloratidine	9.0	7.5[d,h]	NV[g]
Diazepam	9.7	6.9[d]	6.8
Diclofenac	12.4	8.9[i,j]	16.0
Famotidine	8.3	7.8[d,f]	7.4
Indomethacin	9.0	9.3[e,f]	9.5
Irbesartan	9.8	8.8[e]	8.8
Metoprolol	7.8	6.6[e]	7.9
Nadolol	8.4	7.4[d,f]	6.7
Nifedipine	8.8	9.0[d]	8.8
Pirenzepine	9.6	8.0[e]	7.2
Prazosin	8.3	7.4[d]	7.8
Propranolol	8.0	7.7[e]	8.4
Ranitidine	8.7	7.3[e]	7.7
Ritonavir	11.6	8.6[k]	8.6

[a]threshold pX calculated from *in silico* values as in Table 4; [b]threshold pX calculated from *in vitro* values using the same procedure as in Table 4 ($CLint_{HH}$ and fu values from ref [10], P_{app} values from [9] and [14] as specified, and in-house values for fu_{inc}; [c]threshold pX calculated from *in vivo* values [9] using the same procedure as in Table 4; [d]P_{app} value from ref [9]; [e]P_{app} value from ref [14]; [f]fu_{inc} value from in-house *in silico* model; [g]no value (not enough experimental data available for calculation); [h]in-house value for $CLint_{HH}$; [i]fu value from ref [17]; [j]P_{app} from in-house *in silico* model; [k]in-house value for P_{app}.

Most threshold pX values can be found within the range of 7-9, with a couple of values, especially *in vivo* derived, below 7 and others, most often *in silico* derived, above 9. Only one of the compounds for which human *in vivo* PK data was available showed an *in vivo* derived threshold pX values above 10, whereas all *in vitro* derived threshold pX values were actually below 10. The one compound with a higher *in vivo* data derived value, diclofenac, was earlier recognized as not being properly described by the one compartment PK model employed here [9]. For about a third of the compounds all three values are very close within about one log unit, e.g., carvedilol or nitrendipine. Other compounds have a somewhat higher spread, most often the *in vitro* value closer to the *in vivo* derived value, e.g., acebutolol or betaxolol. For a few compounds the *in silico* derived value is clearly differing from the other two, e.g., bisoprolol, diazepam and ritonavir. To better understand why we find the *in silico* predictions differ, we investigated how well the *in silico* and *in vitro* data can predict human *in vivo* properties (see Figures 2-4). Note that cyclosporine and desloratidine were excluded from this analysis, since there was not enough *in vitro* data for the former and no *in vivo* data for the latter compound.
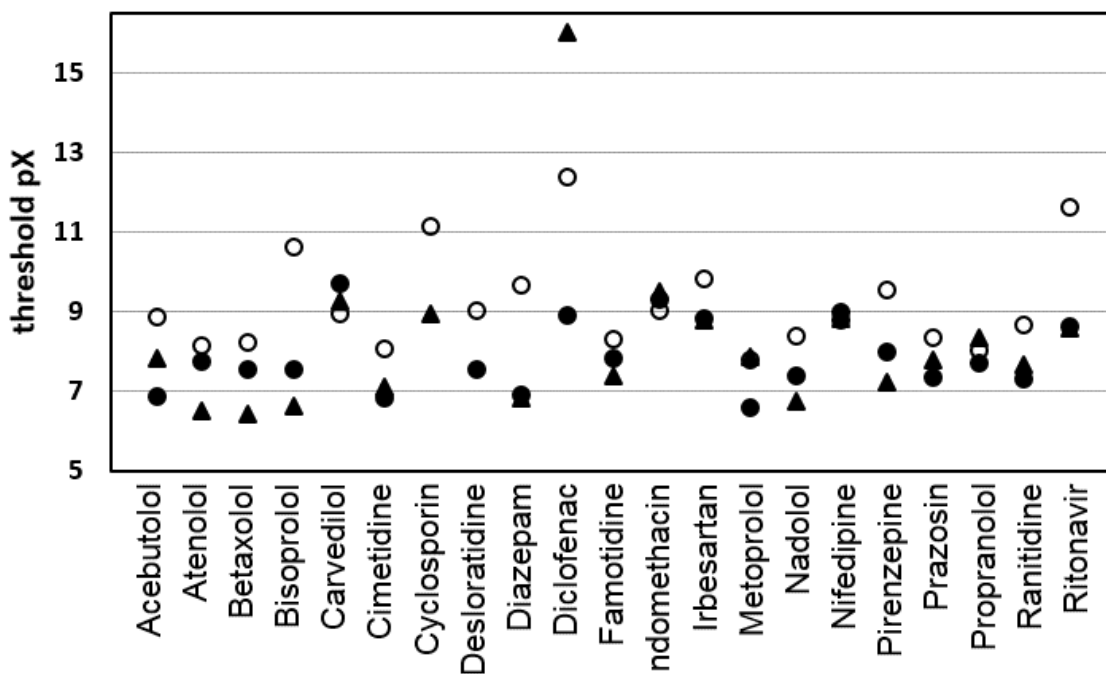
Figure 1. Comparison of threshold pX values calculated from purely *in silico* (open circles), *in vitro* (black circles) or *in vivo* (black triangles) data as described in text.

Figure 2 shows that *in silico* data tends to overpredict human clearance. The two compounds on the left, diazepam and ritonavir, are obvious extremes, and their higher estimate of the threshold potency value are most likely related to the clearance misprediction. Note that clearance predictions from *in vitro* are clearly closer to the line of unity. However, we find two compounds, diclofenac and metoprolol, to be under predicted by more than 3-fold using the present scaling approach. This can also be seen in a more optimistic threshold pX value when compared to *in vivo* (see Figure 1). Overall, about two thirds of the compounds have *in silico* predicted clearance values within 3-fold from the experimental *in vivo* values, and all but two compounds are within 3-fold utilizing *in vitro* values.
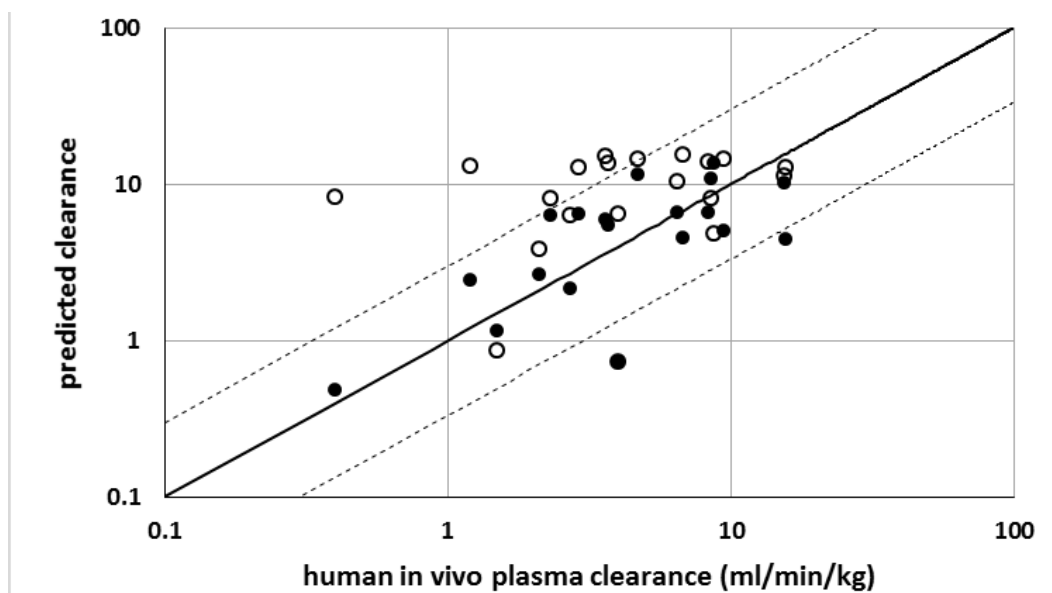
Figure 2. Human *in vivo* plasma clearance plotted against predicted clearance based purely on in silico property calculations (open circles) and on *in vitro* data for human hepatocyte *CLint*, human protein binding and hepatocyte incubation binding (black circles). Full line indicates unity, dashed lines 3-fold limits.

Figure 3. Human *in vivo* plasma half-life plotted against predicted half-life based upon *in silico* volume of distribution and clearance calculated from *in silico* property calculations (open circles) or clearance calculated from *in vitro* data (black circles). Full line indicates unity, dashed lines 3-fold limits.

Also the underprediction of half-life from *in silico* data is mostly due to the clearance prediction, while half-life prediction from *in vitro*, using the same *in silico* derived value for distribution volume, is clearly enhanced.

Human fraction absorbed seems to be the most difficult parameter to predict correctly (see Figure 4). While both *in silico* and *in vitro* predictions identify the two compounds with F_{abs} below 0.4, compounds with F_{abs} between 0.4 and 0.6 were only recognized in two or three cases from *in silico* and *in vitro* data, respectively. However, this failure is not necessarily a concern in the present case, since the relationship between F_{abs} and plasma concentrations is assumed to be linear and a 2-fold change in fraction absorbed will lead to a 2-fold change in plasma concentration and thereby only to a 0.3 log units change for the threshold pX. Actually, Page [10] suggested, that estimation of F_{abs} as either 1 (acids) or 0.5 (all other ion classes) is sufficiently accurate for early D2M predictions.
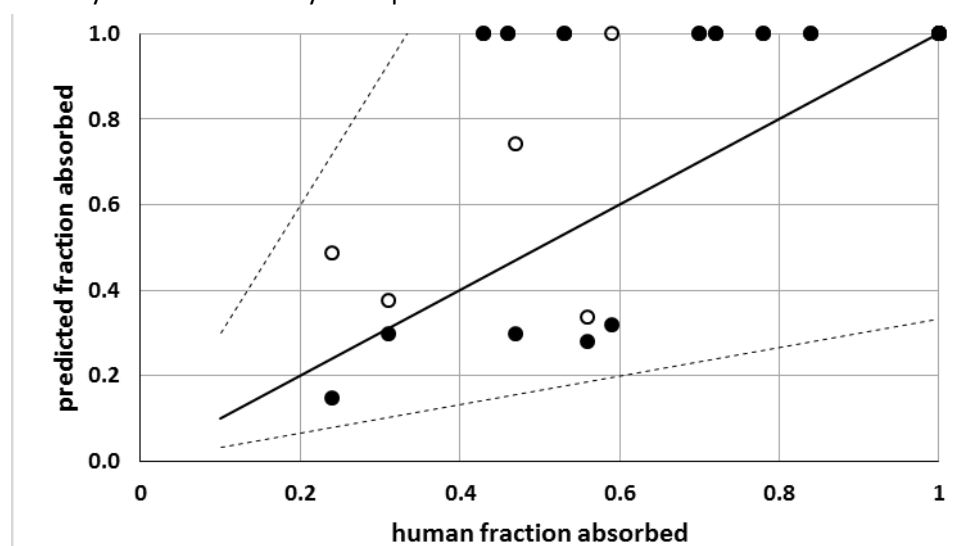

Figure 4. Human fraction absorbed plotted against predicted fraction absorbed from *in silico* Caco permeability prediction (black circles) or *in vitro* Caco permeability data (grey circles). Full line indicates unity, dashed lines 3-fold limits.

Influence of basic parameters – intrinsic clearance, fraction unbound in the hepatocyte incubation, plasma protein binding and volume of distribution

Using pharmacokinetic equations assuming a one compartment model both dose and fraction absorbed are linearly related to the derived plasma concentrations and it is easily understood how either of these parameters will influence the threshold pIC50 calculation. The remaining four input parameters, $CLint_{HH}$, fu_{inc}, fu, and V, on the other hand are more intricately interlinked, especially since the first three also are included in the scaling approach converting *in vitro* (*in silico*) *CLint* to *in vivo* clearance using the well-stirred liver model (equations 5 and 9). In order to better understand how these parameters influence the outcome, threshold pIC50 values for hypothetical compounds with hepatocyte *CLint* values varying from 1 to 200 µl/min/10^6 cells, values for fraction unbound in the incubation varying from 0.01 to 0.7, values for protein binding (*fu*) varying from 0.001 to 0.3, and volume of distribution varying from 0.2 to 3 L/kg were calculated (see Figure 5).

As expected, lower intrinsic clearance as measured in hepatocytes leads to a lower threshold pIC50 estimate. However, the sensitivity of this correlation varies both with the fu_{inc}/fu_b ratio and the volume of distribution: lower volume of distribution leads to higher sensitivity towards intrinsic clearance. Additionally, higher fu_{inc}/fu_b ratios lead to lower sensitivity of the intrinsic clearance as well as of volume of distribution (see Figure 5, panel in upper right corner). Lower *V* leads to higher threshold pIC50 values and higher fu_{inc}/fu_b ratios essentially lead to lower threshold pIC50s, at least as long as potency in blood assays, i.e., defined from total concentrations, is considered. The threshold pIC50 shifts up by three log units in the uppermost panel row (*fu* = 0.001) when free plasma concentrations are taken into account, whereas the threshold pIC50 in the lower panel row (*fu* = 0.3) shifts only by about 0.5 log units.

Sensitivity analysis

Sensitivity analysis is a useful tool for determining which parameters are important. Global sensitivity analysis is used to explore the entire parameter space, considering physiologically relevant values to determine the relative importance of each of them. Here it was shown that volume of distribution is the most influential parameter by far (see Figure 6).

Local sensitivity analysis, on the other hand, can define how sensitive the calculation is to a parameter, when another parameter is scanned over its defined range. Here, we checked the influence of the remaining parameters for different values of volume of distribution (see Figure 7). $CLint_{HH}$ and *fu* have overlapping influence, considering a 10 % change, whereas fu_{inc} is exacty opposite. Since *in vivo CLint* is multiplied by *fu*, when calculating Cl_b using the well stirred model (Eq. 5) and is calculated from $CLint_{HH}$ divided by fu_{inc} (eqs 9 and 9a) this is not really surprising. Note that the influence of the parameters is highest at low volume of distribution values, decreases quickly and does not change further as soon as *V* reaches a value of 2 L/kg. Considering typical volume ranges for acids, bases and neutrals with median values of 0.2 L/kg, 2.9 L/kg and 1.3 L/kg, respectively [10, 17], as indicated in Figure 7, it is clear that especially for acids a small change in either hepatocyte *CLint* or binding properties can make a big difference for the prediction of threshold pIC50 determined from C_{min}.
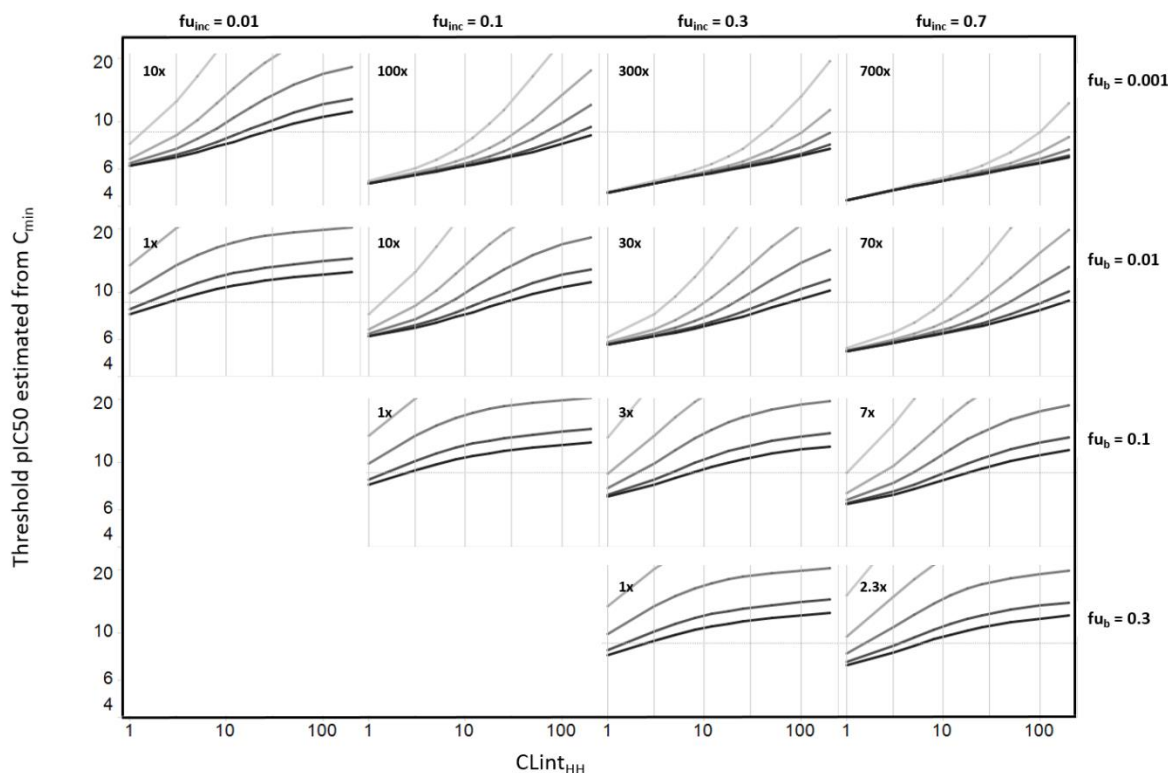
Figure 5. Threshold pIC50 estimated from minimum plasma concentration for hypothetical compounds as calculated from human hepatocyte *CLint* values (1-200 µl/min/10^6 cells, x-axis), fraction unbound in the incubation (fu_{inc} = 0.01 – 0.7; panel columns), fraction unbound in blood (fu_b = 0.001 – 0.3; panel rows), and volume of distribution (V = 3 L/kg, black, V = 2 L/kg, dark grey, V = 1 L/kg, medium grey, V = 0.4 L/kg, light grey, V = 0.2 L/kg, very light grey). Values in the corner of each panel indicate the fu_{inc}/fu_b ratio. Lower left panels left empty since compounds are usually less bound in the hepatocyte incubation than in plasma.

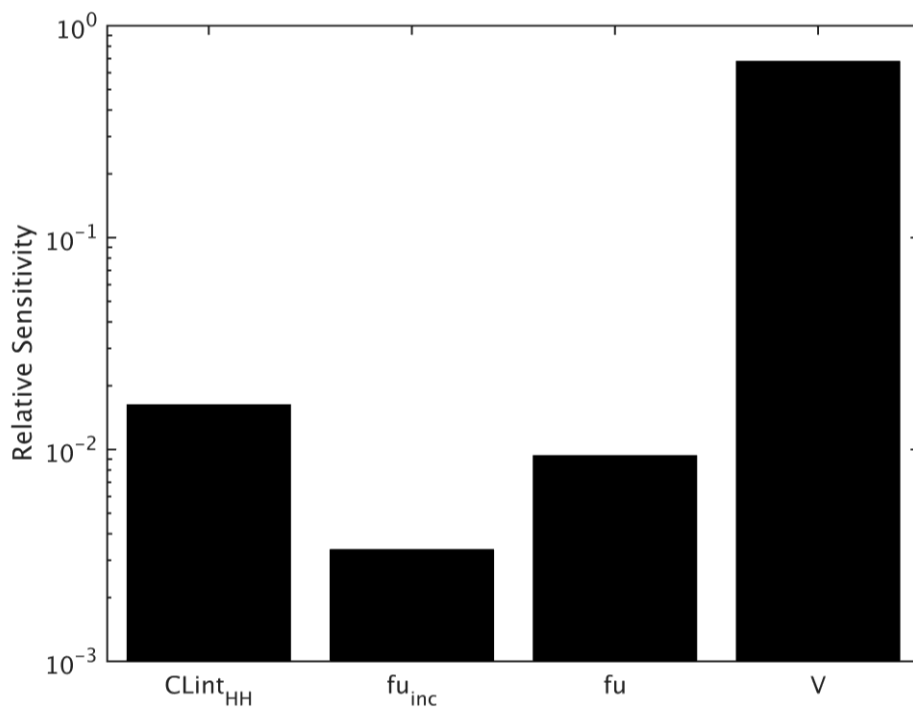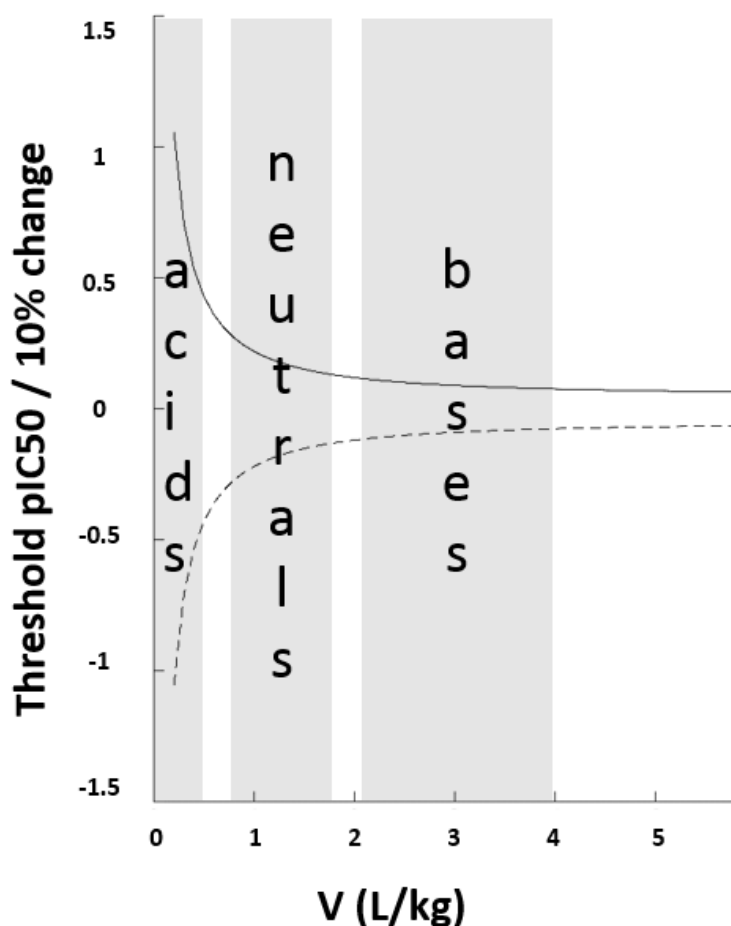**Figure 6.** Global sensitivity towards the four parameters to determine threshold pIC50.

Figure 7. Sensitivity of remaining three parameters for different V values (*CLint*$_{HH}$ and fu full line, *fu*$_{inc}$ dashed line; *CLint* and *fu* have exactly same influence and are on top of each other, *fu*$_{inc}$ has opposite effect), areas of typical distribution volume values for acids, bases and neutrals indicated.

Influence of in silico model quality

The quality of the *in silico* models needs to be considered for the present analysis. Here we use AstraZeneca's in-house models for the five parameters, two of which have rather extensive training sets with more than 10,000 data points each, gathered over a long period of time (>10 years), whereas two, *fu*$_{inc}$ and *P*$_{app}$, have intermediate training set size, and the fifth, volume of distribution, is based on literature data and uses only about 500 compounds. For the four models employing in-house data, about 70-75 % of the compounds within a temporal test set, i.e., compounds that have not been available when the model was built, were found to be within 3-fold of the experimental value. These results are rather encouraging when considering that the experimental variability in general is assumed to be about 2-fold [38]. The models are being regularly updated, as it was shown earlier that continuous updates of models enables new chemistry to be well represented and, thus, likely better predicted by the models [39]. It should be emphasized that the predictivity of the models needs to be investigated for the drug discovery project or chemical series in question when this approach is to be applied. Note that, as was shown in Figure 2, clearance prediction are reasonably accurate from *in silico* for most compounds in data set 2 and can be directly verified as soon as *in vitro* data is available for a compound.

The remaining parameter, human distribution volume, is less easily available. The QSAR model used here is based on human literature data and cannot be updated as straightforwardly as models using data from in-house screens. Also, external data is not necessarily relevant to newest in-house chemistry, and model outcome can usually only be verified at later stages since *in vivo* data is required. Nevertheless, the

predictivity was considered good, with an error of prediction to be 0.4 log units (2.5-fold), and scrutinizing literature or databases for new data [40-42] in fairly regular intervals should ensure continued high quality of the model.

Conclusions

The present study suggests the usage of *in silico* ADME predictions to estimate which plasma concentration a new chemical structure may achieve when given orally using a defined dose and dosing regimen. From the plasma concentration a minimum required potency can be deduced, here referred to as threshold potency (or threshold pIC50). While it was postulated, that early dose predictions from *in silico* was not yet possible [10], the idea here is to utilise predictions only to consider a compound's summarized pharmacokinetic properties for ranking compounds within a series early on in the design process. Combining the *in silico* predictions to a physiologically meaningful score, reduces the risk of deselecting a compound just because one of the parameters is outside an arbitrarily chosen limit.

Using the approach for a set of drug discovery compounds, it was shown that the threshold pIC50 was able to define which compounds had a higher or lower chance of success, with the finally selected compound in the former set. It was assumed that the approach is most useful in lead optimisation stage, when the requirements for potency coverage are known.

Additionally, it was shown that the threshold potency values calculated for a set of marketed drugs were in a reasonable range for most of the compounds, when appropriate dosing regimens were considered for each. These results were also compared with the outcome, when experimental *in vitro* or human *in vivo* data were considered as input for the same procedure. *In silico* outcome was in many cases similar to both *in vitro* and *in vivo*, and when there were bigger differences usually the *in vitro* was closer to the *in vivo* outcome. Thus, *in silico* prediction can be easily verified by *in vitro* experiments as soon as a compound is made.

The examples shown here based threshold potency values on oral dosing and minimum concentration, i.e., required coverage over the whole dosing period. Using such a dosing regimen, and based on a one compartment pharmacokinetic model considering immediate absorption, volume of distribution becomes the most important factor. Special care needs to be taken for low volume compounds such as acids. Obviously, the equations can be adapted to other coverage periods or dosing regimen, as required by the specific drug discovery project.

References

[1]   S. Ekins, J. Rose. *In silico* ADME/Tox: the state of the art. *Journal of Molecular Graphics and Modelling* **20** (2002) 305-309.

[2]   H. van de Waterbeemd, E. Gifford. ADMET in silico modelling: towards prediction paradise? *Nature Reviews Drug Discovery* **2** (2003) 192-204.

[3]   M.P. Gleeson, A. Hersey, S. Hannongbua. In-silico ADME models: a general assessment of their utility in drug discovery applications. *Curr Top Med Chem* **11** (2011) 358-381.

[4]   L. Tao, P. Zhang, C. Qin, S.Y. Chen, C. Zhang, Z. Chen, F. Zhu, S.Y. Yang, Y.Q. Wei, Y.Z. Chen. Recent progresses in the exploration of machine learning methods as in-silico ADME prediction tools. *Advanced Drug Delivery Reviews* **86** (2015) 83-100.

[5]   F. Lombardo, E. Gifford, M.Y. Shalaeva. In silico ADME prediction: data, models, facts and myths. *Mini Reviews in Medicinal Chemistry* **3** (2003) 861-875.

[6]   F. Lombardo, P.V. Desai, R. Arimoto, K.E. Desino, H. Fischer, C.E. Keefer, C. Petersson, S. Winiwarter, F. Broccatelli. In Silico Absorption, Distribution, Metabolism, Excretion, and Pharmacokinetics

(ADME-PK): Utility and Best Practices. An Industry Perspective from the International Consortium for Innovation through Quality in Pharmaceutical Development. *Journal of Medicinal Chemistry* **60** (2017) 9097-9113.

[7]   M.D. Segall, A.P. Beresford, J.M.R. Gola, D. Hawksley, M.H. Tarbit. Focus on success: using a probabilistic approach to achieve an optimal balance of compound properties in drug discovery. *Expert Opinion on Drug Metabolism & Toxicology* **2** (2006) 325-337.

[8]   M.D. Segall. Multi-parameter optimization: identifying high quality compounds with a balance of properties. *Current Pharmaceutical Design* **18** (2012) 1292-1310.

[9]   D.F. McGinnity, J. Collington, R.P. Austin, R.J. Riley. Evaluation of human pharmacokinetics, therapeutic dose and exposure predictions using marketed oral drugs. *Current Drug Metabolism* **8** (2007) 463-479.

[10]   K.M. Page. Validation of Early Human Dose Prediction: A Key Metric for Compound Progression in Drug Discovery. *Molecular Pharmaceutics* **13** (2016) 609-620.

[11]   M. Lemurell, J. Ulander, S. Winiwarter, A. Dahlen, O. Davidsson, H. Emtenas, J. Broddefalk, M. Swanson, D. Hovdal, A.T. Plowright, A. Pettersen, M. Ryden-Landergren, J. Barlind, A. Llinas, M. Herslof, T. Drmota, K. Sigfridsson, S. Moses, C. Whatling. Discovery of AZD6642, an inhibitor of 5-lipoxygenase activating protein (FLAP) for the treatment of inflammatory diseases. *Journal of Medicinal Chemistry* **58** (2015) 897-911.

[12]   J.B. Houston. Utility of in vitro drug metabolism data in predicting in vivo metabolic clearance. *Biochemical Pharmacology* **47** (1994) 1469-1479.

[13]   L.Z. Benet, P. Zia-Amirhosseini. Basic principles of pharmacokinetics. *Toxicologic Pathology* **23** (1995) 115-123.

[14]   L. Fredlund, S. Winiwarter, C. Hilgendorf. In Vitro Intrinsic Permeability: A Transporter-Independent Measure of Caco-2 Cell Permeability in Drug Design and Development. *Molecular Pharmaceutics* **14** (2017) 1601-1609.

[15]   A.-K. Sohlenius-Sternbeck, C. Jones, D. Ferguson, B.J. Middleton, D. Projean, E. Floby, J. Bylund, L. Afzelius. Practical use of the regression offset approach for the prediction of in vivo intrinsic clearance from hepatocytes. *Xenobiotica* **42** (2012) 841-853.

[16]   M.P. Gleeson, A. Hersey, S. Hannongbua. In-Silico ADME Models: A General Assessment of their Utility in Drug Discovery Applications. *Current Topics in Medicinal Chemistry* **11** (2011) 358-381.

[17]   R.S. Obach, F. Lombardo, N.J. Waters. Trend analysis of a database of intravenous pharmacokinetic parameters in humans for 670 drug compounds. *Drug Metabolism and Disposition* **36** (2008) 1385-1405.

[18]   L.S. Goodman, A.G. Gilman, The Pharmacological Basis of Therapeutics, 9th ed., McGraw-Hill Publishers: New York, 1996.

[19]   V. Svetnik, A. Liaw, C. Tong, J.C. Culberson, R.P. Sheridan, B.P. Feuston. Random forest: a classification and regression tool for compound classification and QSAR modeling. *Journal of Chemical Information and Computer Sciences* **43** (2003) 1947-1958.

[20]   P. Bruneau. Search for predictive generic model of aqueous solubility using Bayesian neural nets. *Journal of Chemical Information and Computer Sciences* **41** (2001) 1605-1616.

[21]   ACD/Labs software, 2015, Advanced Chemistry Development, Inc., Toronto, Ontario, Canada.

[22]   ClogP, 4.3, Pomona College and BioByte, Inc., Claremont, CA, US.

[23]   F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, J. Vanderplas, A. Passos, D. Cournapeau, M. Brucher, M. Perrot, É. Duchesnay. Scikit-learn: Machine Learning in Python. *Journal of Machine Learning Research* **12** (2011) 2825-2830.

[24]   H. Wan, F. Bergström. High Throughput Screening of Drug Protein Binding in Drug Discovery. *Journal of Liquid Chromatography & Related Techniques* **30** (2007) 681-700.

[25]   N.J. Waters, R. Jones, G. Williams, B. Sohal. Validation of a Rapid Equilibrium Dialysis Approach for the Measurement of Plasma Protein Binding. *Journal of Pharmaceutical Sciences* **97** (2008) 4586.

[26]   M.C. Wenlock, L.A. Carlsson. How experimental errors influence drug metabolism and pharmacokinetic QSAR/QSPR models. *Journal of Chemical Information and Modeling* **55** (2015) 125-134.

[27]   N. Cristianini, J. Shawe-Taylor, An introduction to support Vector Machines: and other kernel-based learning methods, University Press, Cambridge, 2000.

[28]   J.L. Faulon, D.P. Visco, Jr., R.S. Pophale. The signature molecular descriptor. 1. Using extended valence sequences in QSAR and QSPR studies. *Journal of Chemical Information and Computer Sciences* **43** (2003) 707-720.

[29]   V. Vovk, A. Gammerman, G. Shafer, Algorithmic Learning in a Random World, Springer, New York, 2005.

[30]   GenettaSoft, http://www.genettasoft.com/ (2017-10-25).

[31]   A.-K. Sohlenius-Sternbeck, L. Afzelius, P. Prusis, J. Neelissen, J. Hogstraate, J. Johansson, E. Floby, A. Bengtsson, O. Gissberg, J. Sternbeck, C. Petersson. Evaluation of the human prediction of clearance from hepatocytes and microsome intrinsic clearance for 52 drug compounds. *Xenobiotica* **40** (2010) 637-649.

[32]   D.G. Temesi, S. Martin, R. Smith, C. Jones, B. Middleton. High-throupghput metabolic stability studies in drug discovery by orthogonal acceleration time-of-flight (OATOF) with analogue-to-digital signal capture (ADC). *Rapid Communications in Mass Spectometry* **24** (2010) 1730-1736.

[33]   R.P. Austin, P. Barton, S. Mohmed, R.J. Riley. The binding of drugs to hepatocytes and its relationship to physicochemical properties. *Drug Metabolism and Disposition* **33** (2005) 419-425.

[34]   F. Cannavó. Sensitivity analysis for volcanic source modeling quality assessment and model selection. *Computers & Geosciences* **44** (2012) 52-59.

[35]   MATLAB. The MathWorks Inc., Natick, Massachusetts, US, 2014.

[36]   M. Fink, J.J. Batzel, H. Tran. A Respiratory System Model: Parameter Estimation and Sensitivity Analysis. *Cardiovascular Engineering* **8** (2008) 120-134.

[37]   http://uk.mathworks.com/matlabcentral/fileexchange/15235-automatic-differentiation-for-matlab (2017-10-11).

[38]   S. Winiwarter, B. Middleton, B. Jones, P. Courtney, B. Lindmark, K.M. Page, A. Clark, C. Landqvist. Time dependent analysis of assay comparability: a novel approach to understand intra- and inter-site variability over time. *Journal of Computer-Aided Molecular Design* **29** (2015) 795-807.

[39]   A.M. Davis, D.J. Wood. Quantitative structure-activity relationship models that stand the test of time. *Molecular Pharmaceutics* **10** (2013) 1183-1190.

[40]   F. Lombardo, Y. Jing. In Silico Prediction of Volume of Distribution in Humans. Extensive Data Set and the Exploration of Linear and Nonlinear Methods Coupled with Molecular Interaction Fields Descriptors. *Journal of Chemical Information and Modeling* **56** (2016) 2042-2052.

[41]   M. Olah, R. Rad, L. Ostopovici, A. Bora, N. Hadaruga, D. Hadaruga, R. Moldovan, A. Fulias, M. Mractc, T.I. Oprea, WOMBAT and WOMBAT-PK: Bioactivity Databases for Lead and Drug Discovery, Chemical Biology, Wiley-VCH Verlag GmbH2008, pp. 760-786.

[42]   PharmaPendium PK Fact Sheet. Elsevier, https://www.elsevier.com/__data/assets/pdf_file/0009/80694/R_D-Solutions_PharmaPendium_Fcat-Sheet_-Pharmacokinetics_DIGITAL.pdf (2017-10-24).