JPP JOURNAL OF Pharmacy and Pharmacology
ROYAL PHARMACEUTICAL SOCIETY
JPP Journal of Pharmacy And Pharmacology
Research Paper

Comparison of in-vivo and in-silico methods used for prediction of tissue: plasma partition coefficients in rat

Helen Graham[a], Mike Walker[b], Owen Jones[b], James Yates[b], Aleksandra Galetin[a] and Leon Aarons[a]

[a]School of Pharmacy and Pharmaceutical Sciences, University of Manchester, Manchester and [b]AstraZeneca, Alderley Edge, UK

Keywords
in-silico prediction; partition coefficients; pharmacokinetics

Correspondence
Leon Aarons, School of Pharmacy and Pharmaceutical Sciences, University of Manchester, M13 9PL, Manchester, UK.
E-mail: leon.aarons@manchester.ac.uk

Received February 3, 2011
Accepted November 17, 2011

doi: 10.1111/j.2042-7158.2011.01429.x

Abstract

Objectives To use methods from the literature to predict rat tissue:plasma partition coefficients (*K*ps) and volume of distribution values. Determine which model provides the most accurate predictions to increase confidence in the use of predicted pharmacokinetic parameters in physiologically based pharmacokinetic modelling.

Methods Six models were used to predict *K*ps and four to predict V_{ss} for a dataset of 81 compounds in 11 rat tissues, and the predictions were compared with experimentally derived values.

Key findings *K*p predictions made by the Rodgers *et al*. model were the most accurate, with 77% within threefold of experimental values. The Poulin & Theil model was the most accurate for the prediction of V_{ss}, with 87% of predictions within threefold.

Conclusions This study has shown that in-silico models available in the literature can be used to accurately predict *K*p and V_{ss} in rat. The Rodgers *et al*. model has been shown to provide the most accurate *K*p predictions, with consistent accuracy across all drug classes and tissues. It was also the most accurate V_{ss} predictor when no in-vivo data were used as input. However, transporter systems and other mechanisms that are not yet fully understood need to be incorporated into these types of models in the future to further increase their applicability.

Introduction

Prediction of the likely pharmacokinetic and distribution properties of a compound in humans is an important step in the drug development process, as compounds that are unlikely to exhibit the required properties can be discounted at an early stage, thus reducing the costs associated with development of a new compound. With that in mind, physiologically based pharmacokinetic (PBPK) modelling aims to accelerate and reduce the cost of the drug discovery and development process,[1] by reducing the number of in-vivo and in-vitro experiments required.

These PBPK models aim to accurately represent the processes of absorption, distribution, metabolism and excretion (ADME) using a series of differential equations to mimic the physiological conditions experienced within the body by that drug.[2] These models take into account physiological and biochemical parameters to create a multi-compartment model, with each compartment representing one or more organs of the body. Tissue:plasma partition coefficient (*K*p) values are essential for the development of a PBPK model, as they help to describe the distribution of a drug within the body by defining the ratio of compound concentration between plasma and tissue at equilibrium,[3] and as such they can also be used to predict the volume of distribution at steady state (V_{ss}).[4] Although in-vivo measures of tissue concentration can provide accurate estimates of tissue:plasma partition coefficients, performing these experiments requires the use of significant numbers of animals and resources in processing and analysing the samples. Both are prohibitive to routine data collection in drug discovery where large numbers of compounds need to be tested. Therefore, the development of a method for predicting partition coefficients from in-vitro and in-silico data has become important for the future of PBPK modelling. These methods mean that *K*ps can be predicted from more easily accessible parameters, which may already be available in the literature or can be easily obtained experimentally.

Early *K*p prediction methods in the literature calculated *K*ps by using the octanol:water partition coefficient of the drug of interest to describe distribution, but Poulin and Krishnan[5] discovered that partitioning in vegetable oil was

a better way to represent the solubility of an organic drug in the neutral lipids present in adipose tissue. Other studies have used quantitative structure–property relationships as a basis for their predictions.[6,7] More recently, the work of Poulin and Krishnan has been expanded to take into account the lipid, phospholipid and water content of tissues, and the solubility of a compound in these three phases, to develop a mechanistic algorithm for the prediction of *K*p values.[8] Models designed to improve upon this work were also developed, such as the Rodgers *et al*. model[9] that aimed to improve the prediction of moderate-to-strong bases, and the Berezhkovskiy model.[10] Other recent models have taken a more empirical approach, and based their predictions on in-vivo data, such as volume of distribution,[11] or have used measured in-vivo partition coefficients in one tissue (e.g. muscle) as a predictor for the *K*p values of all other tissues.[12]

Previous studies have compared the ability of numerous *K*p prediction methods to predict rat and human V_{ss}.[13,14] The Jones *et al*. study undertook a comprehensive analysis of the ability of 24 methods to predict V_{ss} in human, using a dataset of 18 blinded compounds.[13] However, no comprehensive analysis using a dataset of this size and composed of literature compounds has previously been attempted. Additionally, no previous study has validated the results at the *K*p level as well as determining the accuracy of the V_{ss} predictions in rat.

Therefore, the objective of this study was to use six established methods available in the literature to produce *K*p predictions for a dataset of 81 drug compounds in 11 rat tissues, and to compare these predictions to experimental values in order to analyse their accuracy. Models selected for this study varied in complexity and the amount of experimental data required (details in Materials and Methods). Drug selection in the current analysis was based on availability of input parameters across all methods investigated and the availability of corresponding experimental *K*p data and comprises acidic, basic, neutral and zwitterionic drugs. *K*p predictions made by four of the models were then used to predict V_{ss} values for the same dataset of compounds, and the accuracy of these results analysed by comparison to in-vivo V_{ss} values.

Materials and Methods

The six methods of prediction compared in this study were taken from the literature (Arundel,[11] Berezhkovskiy,[10] Jansson *et al*.,[12] Poulin *et al*.,[8,15] Poulin & Theil[16] and Rodgers *et al*.[9,17]). Three of these models can be categorised as being mechanistic or *in silico*, requiring only tissue composition data and physicochemical drug properties as input (i.e. Poulin *et al*., Berezhkovskiy, and Rodgers *et al*.). The remaining three models take an empirical or in-vivo approach, requiring experimentally derived data along with the physicochemical properties of the compound (i.e. Arundel, Jansson *et al*., and Poulin & Theil).

In some cases, more than one equation has been developed by the author(s) to deal with different tissue types and drug classes. The model by Rodgers *et al*. comprises one equation for the prediction of *K*p for moderate-to-strong bases and another for all other drug classes,[9,17] whereas the Poulin *et al*. model uses one equation for all tissues except adipose, for which a second equation was generated.[8,15] For the purpose of this study, these equations were grouped together under one heading (e.g. 'Rodgers *et al*. method' refers generally to all equations by this author).

The Rodgers *et al*. model and the Poulin & Theil model predict values for tissue:plasma water partition coefficients (i.e. *K*pu rather than *K*p). Therefore, the values predicted by these two methods were expressed as *K*p by using the relationship shown in Equation 1 using fu_p (fraction unbound in plasma) values taken from the literature (Table S1).

$$Kp = Kpu^{*}\, fu_p \tag{1}$$

In-vivo models

Arundel model

The Arundel model[11] is classed as an in-vivo model because it requires in-vivo V_{ss} as an input parameter. This multi-compartmental model approach was based upon the principle that organs and tissues can be 'lumped' together as long as they occupy parallel positions in the system structure and have similar time constants.[18] Therefore this model deals with six lumped compartments: (1) lung, (2) brain, heart and kidney, (3) gut and spleen, (4) liver, (5) muscle and (6) adipose. The parameter known as the tissue rate constant (Kt_i), or rate of disappearance from the tissue of a drug *i*, is used to characterise these compartments.

Arundel found that the product of $(Kt \times V_{ss})_j$ for the lumped compartments remained fairly constant for all tissues (except adipose), and therefore based on the steady-state volume of distribution of a compound (V_{ss}) its Kt_i values can be calculated using Equation 2:

$$Kt_i = \frac{(Kt \times V_{ss})_j}{(V_{ss})_i} \tag{2}$$

Using this equation and the relationship shown in Equation 3 (where *Per* represents the perfusion rate of the tissue, normalised for tissue size), partition coefficients for each lumped compartment can be calculated by integration of the Kt_i values obtained.

$$Kp = \frac{Per_i}{Kt_i} \tag{3}$$

The *K*p value for each lumped compartment is in effect an 'average' *K*p of all the tissues it represents, but the model can be expanded to calculate the individual *K*p values as shown in Equation 4, with *Qt* representing the blood flow rate of the tissue, and Vt_i representing the volume of the tissue.

$$Kt_i = \frac{Qt_i}{Vt_i \times Kp} \quad (4)$$

For adipose tissue, a different method was required, whereby Arundel proposed that the partition coefficient could be estimated using $\log D_{7.4}$ as an input parameter to Equation 5:

$$\log Kp = -0.6 + 0.8 * \log D_{7.4} \quad (5)$$

Jansson et al. *model*

The Jansson *et al.* model[12] is an empirical prediction method based on a combination of a measured volume of distribution, and a lipophilicity descriptor of the compound. Furthermore, linear regression analysis is used to predict *K*p values for all other tissues from the in-vivo *K*p value in muscle. This method of linear regression was first described by Björkman[4] and is detailed in Equation 6:

$$Kp_{tissue} = \text{slope}_{tissue} \times Kp_{muscle} + \text{intercept}_{tissue} \quad (6)$$

Jansson *et al.* incorporated this regression into their model while also using a lipophilicity descriptor ($\log X_{drug}$) for certain tissues where it was deemed necessary. The $\log X_{drug}$ parameter can be either logP, $\log D_{7.4}$ (logP adjusted for ionisation at pH 7.4), or $\log K_{7.4}$ (vegetable oil:water partitioning adjusted for ionization at pH 7.4). The selection of descriptor to use is based upon the use of an *F*-test, and the descriptor with the lowest sum of squared residuals is chosen. Using this and values for Kp_{muscle} taken from the literature, two equations were developed (Equations 7 and 8) for the prediction of partition coefficients (where b_{tissue} is the slope for $\log X_{drug}$):

$$Kp_{tissue} = 10^{slope_{tissue} \times \log Kp_{muscle} + \text{intercept}_{tissue}} \quad (7)$$

$$Kp_{tissue} = 10^{slope_{tissue} \times \log Kp_{muscle} + b_{tissue} \times \log X_{drug} + \text{intercept}_{tissue}} \quad (8)$$

Poulin & Theil model

The Poulin & Theil model[16] is similar to the Jansson *et al.* model described above in that it uses linear regression to define the relationship between the in-vivo muscle *K*pu value and the *K*pu of all other tissues. Equation 9 describes the relationship, where *m* is the slope and $\log(b)$ the intercept:

$$\log(Kpu) = m \log(Kpu_{muscle}) + \log(b) \quad (9)$$

For adipose tissue, Poulin & Theil suggested a different approach, using an adjusted skin *K*pu value to predict adipose *K*pu.

In-silico models

Poulin et al. *model*

The main aim of the equations derived by Poulin *et al.*[8,15] was to incorporate two factors into the prediction of tissue-:plasma partition coefficients – the solubility of a drug in lipids, and the binding of a drug to macromolecules. In total, Poulin *et al.* devised three equations to predict *K*p values for all drug types. The first equation does not apply to drugs that reside predominantly in the interstitial space of tissues, and instead applies only to those drugs that assume a homogeneous distribution (Equation 10), where $K_{vo:w}$ is the distribution in vegetable oil:water, V_n is the neutral lipid content, V_{ph} is the phospholipid content and V_w is the water content.

$$Kp = \frac{\left[K_{vo:w}\left(V_{nt} + 0.3V_{pht}\right)\right] \times \left[\left(V_{wt} + 0.7V_{pht}\right)\right]}{\left[K_{vo:w}\left(V_{np} + 0.3V_{php}\right)\right] \times \left[\left(V_{wp} + 0.7V_{php}\right)\right]} \cdot \frac{fu_p}{fu_t} \quad (10)$$

A further equation devised by Poulin *et al.* only applies to those drugs that reside predominantly in the interstitial space (Equation 11), where F_t represents the fractional content of interstitial space in tissue and F_p represents the fractional content of interstitial space in plasma.

$$Kp = \left(F_t / F_p\right) \cdot \left(fu_p / fu_t\right) \quad (11)$$

This equation is derived from the observation that experimentally obtained *K*p values for these types of drugs are approximately equal to the ratio between the interstitial volumes of tissues and plasma.

A third equation was generated to predict *K*p values for adipose tissue (Equation 12).

$$Kp = \frac{\left[K^{\star}_{vo:w}\left(V_{nt} + 0.3V_{pht}\right)\right] \times \left[\left(V_{wt} + 0.7V_{pht}\right)\right]}{\left[K^{\star}_{vo:w}\left(V_{np} + 0.3V_{php}\right)\right] \times \left[\left(V_{wp} + 0.7V_{php}\right)\right]} \cdot \frac{fu_p}{1} \quad (12)$$

This equation differed from that designed for non-adipose tissues in that fu_t is now set to 1 in order to remove the effect of macromolecular binding, and $K_{vo:w}$ is replaced with $K^{\star}_{vo:w}$ (where $K_{vo:w}$ is equivalent to the distribution of non-ionised species in oily and aqueous phases and $K^{\star}_{vo:w}$ is equivalent to the distribution between non-ionised species and ionised and non-ionised species in oily and aqueous phases, respectively). These changes were made to reflect the markedly different behaviour exhibited by compounds in adipose tissue when compared with other tissues, in which the effect of macromolecular binding was found to be negligible.

Berezhkovskiy model

Berezhkovskiy described how the Poulin *et al*. model relies on the following relationship (Equation 13):

$$\frac{(V_{wt}+0.7V_{pht})}{(V_{wp}+0.7V_{php})}=1 \quad (13)$$

The Berezhkovskiy model[10] therefore represents a modified version of the Poulin *et al*. model, which does not require the assumption embodied in Equation 13 (Equation 14):

$$Kp=\frac{\left[K_{vo:w}(V_{nt}+0.3V_{pht})+0.7V_{pht}+\frac{V_{wt}}{fu_t}\right]}{\left[K_{vo:w}(V_{np}+0.3V_{php})+0.7V_{php}+\frac{V_{wp}}{fu_p}\right]} \quad (14)$$

Therefore, tissue binding is being considered only in the water fraction.

Rodgers et al. *model*

Two models were devised by Rodgers *et al*. – one to predict tissue:unbound plasma water partition coefficients (*K*pu) for moderate-to-strong bases,[9] and another to predict *K*pu for acids, very weak bases, neutrals and Group 2 zwitterions (i.e. those that do not have a pKa $\geq$ 7).[17]

Equation 15 (where EW represents extracellular water, IW represents intracellular water, NL is neutral lipid, NP is neutral phospholipids, AP^- is acidic phospholipid, *K*a is the association constant of basic compounds with AP, p is plasma, and *P* is the octanol:water partition coefficient for all tissues except adipose, for which *P* represents the vegetable oil:water partition coefficient) accommodates the electrostatic interactions that form between basic drugs and acidic tissue phospholipids. This equation incorporates the partitioning of drug into neutral lipids and phospholipids, and also the dissolution of the drug into tissue water.

$$Kpu=\left[f_{EW}+\left(\frac{1+10^{pKa-pH_{IW}}}{1+10^{pKa-pH_p}}\cdot f_{IW}\right)+\left(\frac{Ka\cdot[AP^-]_T\cdot 10^{pKa-pH_{IW}}}{1+10^{pKa-pH_p}}\right)+\left(\frac{(P\cdot f_{NL}+((0.3P+0.7)\cdot f_{NP}))}{1+10^{pKa-pH_p}}\right)\right] \quad (15)$$

The second equation (Equation 16, where PR represents protein) also incorporates partitioning into neutral lipids and phospholipids, and the dissolution of a drug into the tissue water. Furthermore, it incorporates associations with extracellular proteins.

$$Kpu=\frac{X\cdot f_{IW}}{Y}+f_{EW}+\left(\frac{P\cdot f_{NL}+(0.3P+0.7)\cdot f_{NP}}{Y}\right)+\left[\left(\frac{1}{fu}-1-\left(\frac{P\cdot f_{NL,P}+(0.3P+0.7)\cdot f_{NP,P}}{Y}\right)\right)\cdot\frac{[PR]_T}{[PR]_P}\right] \quad (16)$$

V_{ss} calculation

V_{ss} (volume of distribution at steady state) values were predicted using predicted *K*p values from four of the models – those authored by Berezhkovskiy, Poulin *et al*., Rodgers *et al*. and Poulin & Theil. The Arundel and Jansson *et al*. models could not be used to predict V_{ss}, as they use in-vivo V_{ss} as an input parameter.

V_{ss} was calculated using Equation 17:

$$V_{ss}=\sum(Kp*V_t)+V_p \quad (17)$$

where V_t represents tissue volume and V_p is plasma volume.

For comparison purposes, this equation was also used to predict V_{ss} from experimentally derived *K*p values.

Dataset selection

*K*p values were predicted for 81 drugs – 21 acids, 47 bases, 5 neutrals and 9 zwitterions. The logP values of this dataset ranged from −4.51 to 6.3 (for pyridostigmine and *trans*-retinoic acid, respectively) and pKa values ranged from −1.6 to 10.4 (for pyridostigmine and caffeine, respectively). These drugs were chosen specifically to ensure that experimentally derived *K*p data were available for at least three tissues for each of the compounds in the dataset (73 of the compounds had *K*ps for five or more tissues available). Data were gathered from Jansson *et al*.,[12] Rodgers *et al*.,[9,17] and Poulin *et al*.,[15] and all experimental *K*p values quoted were selected by these authors as they were found to plausibly represent steady-state distribution or pseudo equilibrium. In-vivo V_{ss} values for all drugs in the dataset were collated from the literature. All *K*p and V_{ss} values were determined in rat.

Compound specific input parameters

The compound specific input parameters for this dataset were taken from the literature and are summarised in Table S1. The only exception to this is the vegetable oil : water partition coefficient, $\log P_{vo:w}$, which was calculated using the following relationship (Equation 18):[19]

$$\log P_{vo:w}=1.115*\log P_{o:w}-1.35 \quad (18)$$

Tissue specific input parameters

*K*p values were predicted for adipose, bone, brain, gut, heart, kidney, liver, lung, muscle, skin and spleen tissue, with the

exception of the Poulin & Theil (2009) model which does not predict muscle *K*p, as it uses the experimentally-derived in-vivo muscle *K*p as an input parameter. These 11 tissues were then also used to predict V_{ss} values. The tissue specific input parameters and tissue volumes for rat were taken from the literature, and are summarized in Table S2.

Statistics

Accuracy of the *K*p and V_{ss} predictions was assessed by determining the percentage of predictions which fell within < threefold, three to fivefold, and > fivefold of observed values. The following statistical analyses were also performed:

Accuracy was assessed using afe (average fold error; Equation 19) and aafe (absolute average fold error; Equation 20), where afe is used to assess to what extent a method over-predicts or under-predicts the experimentally determined values, and aafe gives the absolute value of the error.

$$\text{afe} = 10^{\left(\frac{1}{n}\sum \log(PRED_i/OBS_i)\right)} \tag{19}$$

$$\text{aafe} = 10^{\left(\frac{1}{n}\sum \log(|PRED,i/OBS,i|)\right)} \tag{20}$$

where $PRED_i$ and OBS_i refer to the predicted and observed *K*p value for the *i*th compound respectively.

Precision was assessed using rmse (root mean squared error; Equation 21):

$$\text{rmse} = \sqrt{\left(\frac{1}{n}\sum (PRED_i - OBS_i)^2\right)} \tag{21}$$

The correlation concordance coefficient (ccc) was also calculated as follows (Equation 22):

$$ccc = \frac{2s_{xy}}{s_x^2 + s_y^2 + (\bar{x} - \bar{y})^2} \tag{22}$$

s and s^2 are the covariance and the variance, as defined below (Equations 23 and24):

$$s_{xy} = \frac{1}{n}\sum (x_i - \bar{x})(y_i - \bar{y}) \tag{23}$$

$$s_x^2 = \frac{1}{n}\sum_{i=1}^{n} (x_i - \bar{x})^2, s_y^2 = \frac{1}{n}\sum_{i=1}^{n} (y_i - \bar{y})^2 \tag{24}$$

where x_i is the predicted *K*p value of the *i*th compound, y_i is the observed value, $\bar{x}$ is the average of the predicted values, and $\bar{y}$ the average of the observed values.

Statistically significant differences between pairs of models were assessed using a chi square test (Equation 25) to compare the number of predictions that fall within, and outside of, threefold of experimental values (level of significance set at $P < 0.05$).

$$\chi^2 = \frac{(a-c)^2}{c} + \frac{(b-d)^2}{d} \tag{25}$$

where *a* and *b* refer to the number of predictions made by model 1 which are within and outside of threefold of experimental values respectively, and *c* and *d* refer to the same values for model 2.

Results

*K*p predictions

Predicted *K*p values for each of the 81 drugs and 11 tissues using all the methods investigated are summarised in Table S3. The model with the highest percentage of *K*p predictions within threefold of experimental values was the Rodgers *et al.* with 77.3% (89.4% within fivefold). This model had significantly more predictions within the threefold range than any of the other models ($P < 0.01$). The rank order of remaining methods was Poulin & Theil > Jansson *et al.* > Arundel > Poulin *et al.* > Berezhkovskiy (Figure 1). In the case of the latter two models, < 51% of the predictions were within threefold of the experimental values (Table 1). The Poulin & Theil, Jansson *et al.* and Arundel models all resulted in comparable overall *K*p prediction success (68–71% within threefold). There was no significant difference between the number of predictions within the threefold range for the Arundel and Jansson *et al.* models ($P = 0.96$). For the current dataset only the Poulin *et al.* and Berezhkovskiy models had > 30% of *K*p predictions outside of the fivefold margin.

The afe of the Berezhkovskiy, Poulin *et al.* and Poulin & Theil models is < 1, indicating that all of these models exhibit a general tendency to under-predict the *K*p values across all drug classes and tissue types.

Assessment of prediction accuracy by drug class

For acidic compounds, the Jansson *et al.* model showed the highest accuracy of *K*p prediction (90.2% of predictions within threefold of experimental values, afe of 1.26). The Rodgers *et al.*, Arundel, Poulin *et al.* and Berezhkovskiy models all performed well for this drug class with more than 70% of predictions within threefold of experimental values (Table 2). The number of predictions within the threefold range made by the Jansson *et al.* model were significantly higher than the number of predictions made by all of the other models ($P < 0.01$) except the Rodgers *et al.* model. The Poulin & Theil model had the lowest percentage of predictions within the threefold range with 55.8%.

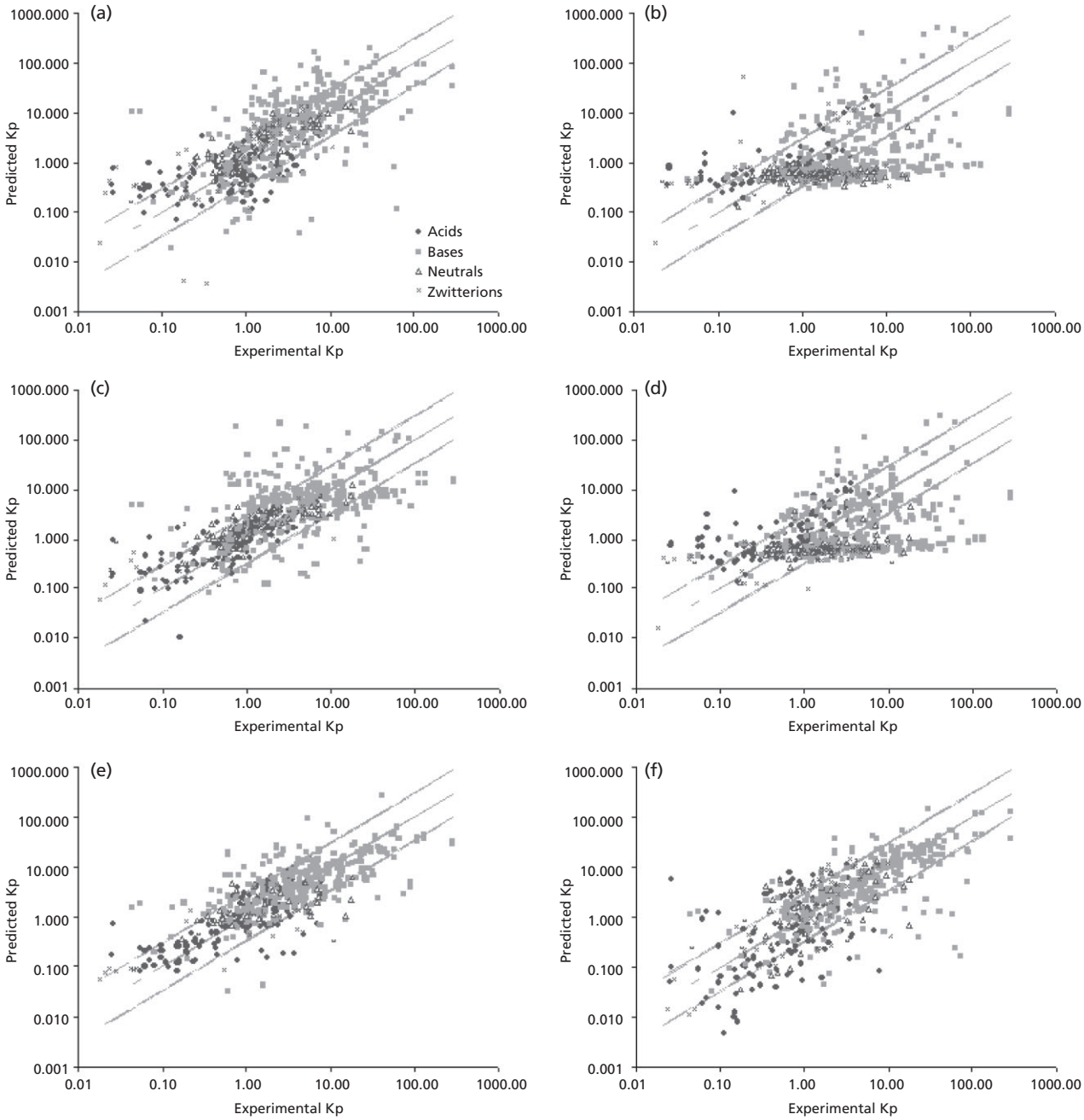

Figure 1 Relationship between predicted and experimentally derived *Kp* values for 81 compounds using the Arundel model (a), the Berezhkovskiy model (b), the Jansson *et al*. model (c), the Poulin *et al*. model (d), the Rodgers *et al*. model (e) and the Poulin & Theil model (f). Straight lines indicate line of unity and threefold above and below; diamonds indicate acids; squares indicate bases; triangles indicate neutrals; and crosses indicate zwitterions.

For basic compounds, the Poulin & Theil model performed significantly ($P < 0.05$) better than any other model with 80.4% of compounds within the threefold range, with the Berezhkovskiy model having the lowest percentage of predictions within this range (32.2%; Table 3). Jansson *et al.* produced the most accurate predictions for both neutral compounds (83.8% within threefold; results not significant) and zwitterionic compounds (80.9%; $P < 0.01$).

Predictions for the Arundel model showed the highest consistency across all drug classes, with predictions differing by just 9.7% (from 65.3% for basic compounds to 75.0% for neutrals). In contrast, predictions by the Berezhkovskiy

Table 1 Comparative and statistical assessment of the six models used to predict *K*p using a dataset of 81 compounds (21 acids, 46 bases, 5 neutrals, 9 zwitterions) in 11 rat tissues

Model	% < threefold	% three to fivefold	% > fivefold	afe	aafe	rmse	ccc	*n*	Mean pred : obs ratio	SD
Arundel	68.5	13.6	17.8	1.15	2.62	0.58	0.51	623	2.97	12.94
Berezhkovskiy	46.1	18.5	35.8	0.45	4.20	0.79	0.16	645	2.13	11.68
Jansson *et al*.	68.7	14.8	16.7	1.08	2.58	0.57	0.24	623	3.19	12.72
Poulin *et al*.	50.5	17.8	31.8	0.54	3.78	0.73	0.16	645	1.76	4.33
Rodgers *et al*.	77.3	12.1	10.6	1.01	2.13	0.44	0.42	586	1.76	3.11
Poulin & Theil	70.5	14.5	15.5	0.85	2.41	0.53	0.72	543	1.59	2.33

aafe, absolute average fold-error; afe, average fold-error; ccc, correlation concordance coefficient; obs, observed; pred, predicted; rmse, root mean squared error; SD, standard deviation.

Table 2 Comparative and statistical assessment of the six models used to predict *K*p in rat using 21 acidic compounds

Model	% < threefold	% three to fivefold	% > fivefold	afe	aafe	rmse	ccc	*n*	Mean pred : obs ratio	SD
Arundel	75.0	14.5	10.5	1.03	2.22	0.44	0.51	153	1.84	3.01
Berezhkovskiy	71.7	14.5	13.9	0.99	2.30	0.49	0.41	166	2.40	6.50
Jansson *et al*.	90.2	5.9	4.6	1.26	1.70	0.33	0.79	153	1.86	3.47
Poulin *et al*.	74.1	12.7	13.3	1.47	2.35	0.51	0.34	166	3.39	7.55
Rodgers *et al*.	85.0	10.2	6.0	1.15	1.91	0.39	0.53	167	1.77	3.14
Poulin & Theil	55.8	17.8	28.7	0.80	3.34	0.66	0.20	129	3.57	19.33

aafe, absolute average fold-error; afe, average fold-error; ccc, correlation concordance coefficient; obs, observed; pred, predicted; rmse, root mean squared error; SD, standard deviation.

Table 3 Comparative and statistical assessment of the six models used to predict *K*p in rat using 46 basic compounds

Model	% < threefold	% three to fivefold	% > fivefold	afe	aafe	rmse	ccc	*n*	Mean pred : obs ratio	SD
Arundel	65.3	13.1	21.7	1.22	2.78	0.60	0.37	360	3.73	16.83
Berezhkovskiy	32.2	19.2	48.6	0.31	5.69	0.91	0.13	360	1.61	5.53
Jansson *et al*.	56.1	20.7	23.2	1.01	3.19	0.66	0.11	353	4.27	16.69
Poulin *et al*.	41.1	17.3	41.6	0.39	4.77	0.84	0.12	353	1.35	2.67
Rodgers *et al*.	75.7	13.6	10.7	1.06	2.18	0.47	0.30	337	2.05	3.96
Poulin & Theil	80.1	9.6	10.3	0.83	2.08	0.49	0.56	301	1.45	2.49

aafe, absolute average fold-error; afe, average fold-error; ccc, correlation concordance coefficient; obs, observed; pred, predicted; rmse, root mean squared error; SD, standard deviation.

model differed by nearly 40% across classes (32.2% for basic compounds to 71.7% for acidic compounds).

Assessment of prediction accuracy by tissue

Large differences in accuracy of prediction between different tissues were observed. Across all models, based on accuracy of prediction, the tissues fall into the following order: skin (an average across all models of 75.2% of predictions within threefold of experimental values), heart (72.7%), muscle (69.8%), gut (65.1%), lung (61.7%), liver (61.2%), brain (60.4%), spleen (58.3%), kidney (58%), bone (52.4%) and adipose (51.3%). The highest percentage of predictions within 3-fold of experimental values was generated by the Poulin & Theil model in heart tissue, with 97.1% within 3-fold, an afe of 0.80 and an rmse of 0.21. The lowest percentage of predictions within threefold was made by the Poulin & Theil model in spleen tissue, with just 20.0%. A high afe of 4.45 indicated that this model consistently over-predicted *K*ps for this tissue. Poor prediction accuracy was observed for spleen *K*p regardless of the model used (on average 58% within threefold), with aafe values ranging from 1.78 to 4.45 for the Rodgers *et al.* and Poulin & Theil models, respectively. The Rodgers *et al.* model performed the best for this tissue,

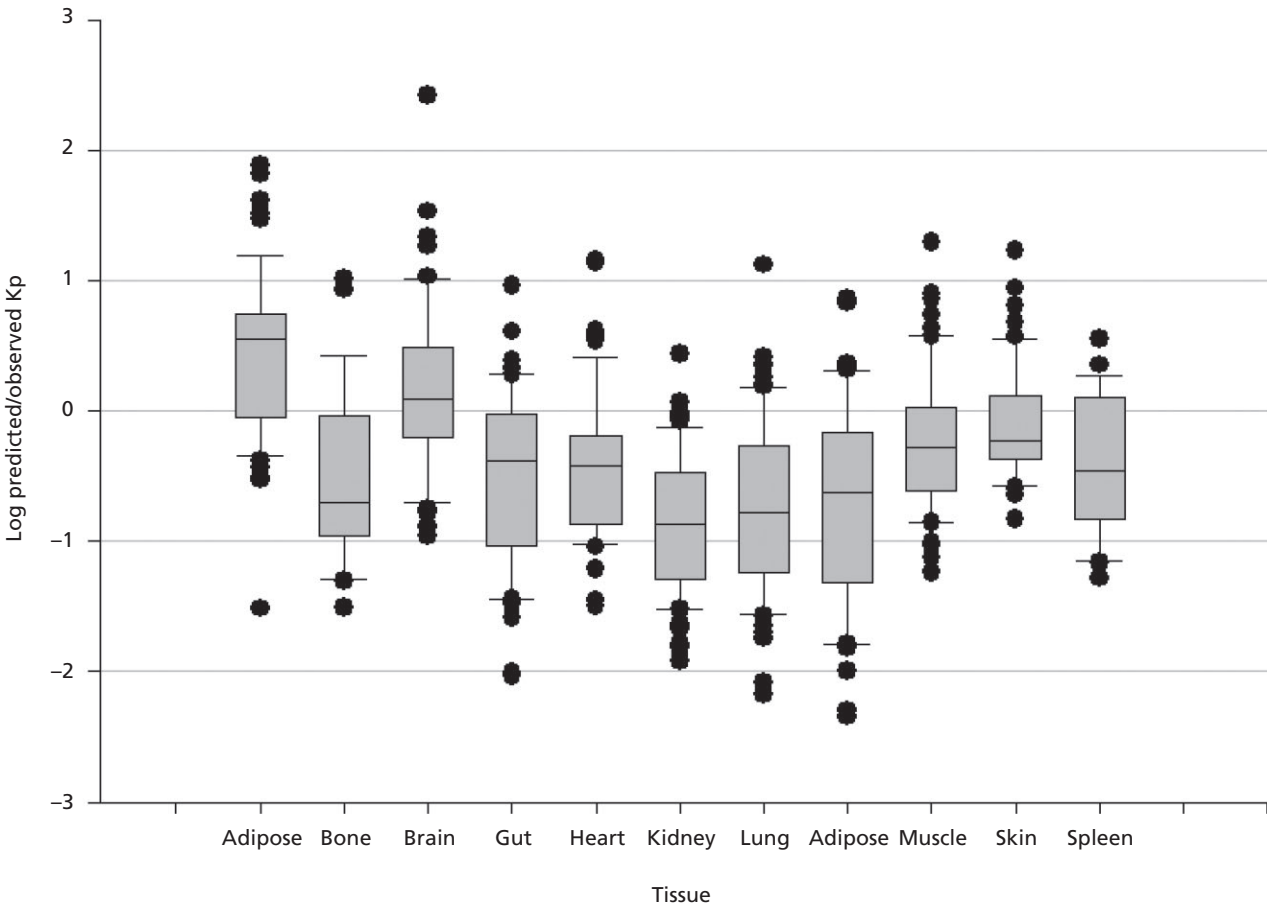

Figure 2 Boxplot of log predicted/observed *K*p values for each tissue as predicted by the Berezhkovskiy model.[10] The boxes indicate the standard deviation; the whiskers represent the range between the 10th and 90th percentile; the horizontal black line represents the median value; and outliers are represented by black circles.

with 81.5% of predictions within the threefold range, followed by the Poulin *et al.* model with 68.2%.

The difference between models is particularly evident for kidney tissue, where the difference between the highest and lowest percentage of predictions within threefold was almost 50% (Berezhkovskiy 26.1% vs Arundel 76.8%). The fold error for this tissue was also highly variable, with aafe values ranging from 1.86 (Poulin & Theil) to 6.08 (Poulin *et al.*). A similar difference was seen for liver tissue, with a difference of over 40% between the most accurate model (72.6%; Rodgers *et al.*) and the least accurate model (31.3%; Berezhkovskiy). In contrast, skin was predicted with high accuracy by all models, ranging from 64.9% (Poulin *et al.* model) to 83.3% (Rodgers *et al.* model). This high accuracy was reflected in the aafe values, ranging from 1.97 (Poulin & Theil) to 2.53 (Poulin *et al.*).

Predictions for lung tissue were particularly inaccurate, with the exception of the Poulin & Theil model (aafe values ranging from 2.27 in the Rodgers *et al.* model to 6.61 in the Berezhkovskiy model). All models under-predicted *K*p values for this tissue, except the Poulin & Theil model (afe of 1.13), which was the most accurate and precise model for this tissue type, with 83.3% of predictions in the threefold range, and an rmse of 0.34.

The Berezhkovskiy model under-predicted *K*p values for all tissues except adipose and brain (Figure 2), which it over-predicted with afe values of 2.83 and 1.38, respectively. It was the only model to show such a strong bias towards either over- or under-prediction. The Rodgers *et al.* model over-predicted *K*p for five out of the eleven tissues studied (adipose, brain, heart, muscle and skin; Figure 3) with afe values ranging from 1.19 for adipose to 1.96 for skin. For the five tissues for which this model under-predicted *K*p, the afe values ranged from 0.55 for kidney to 0.98 for spleen (for the remaining tissue, bone, the afe value is 1.0, indicating that this tissue did not show a bias towards either under-prediction or over-prediction).

The Rodgers *et al.* model was consistently the most accurate across all tissues, with the percentage of predictions within threefold differing by 26.6% across all tissues (from

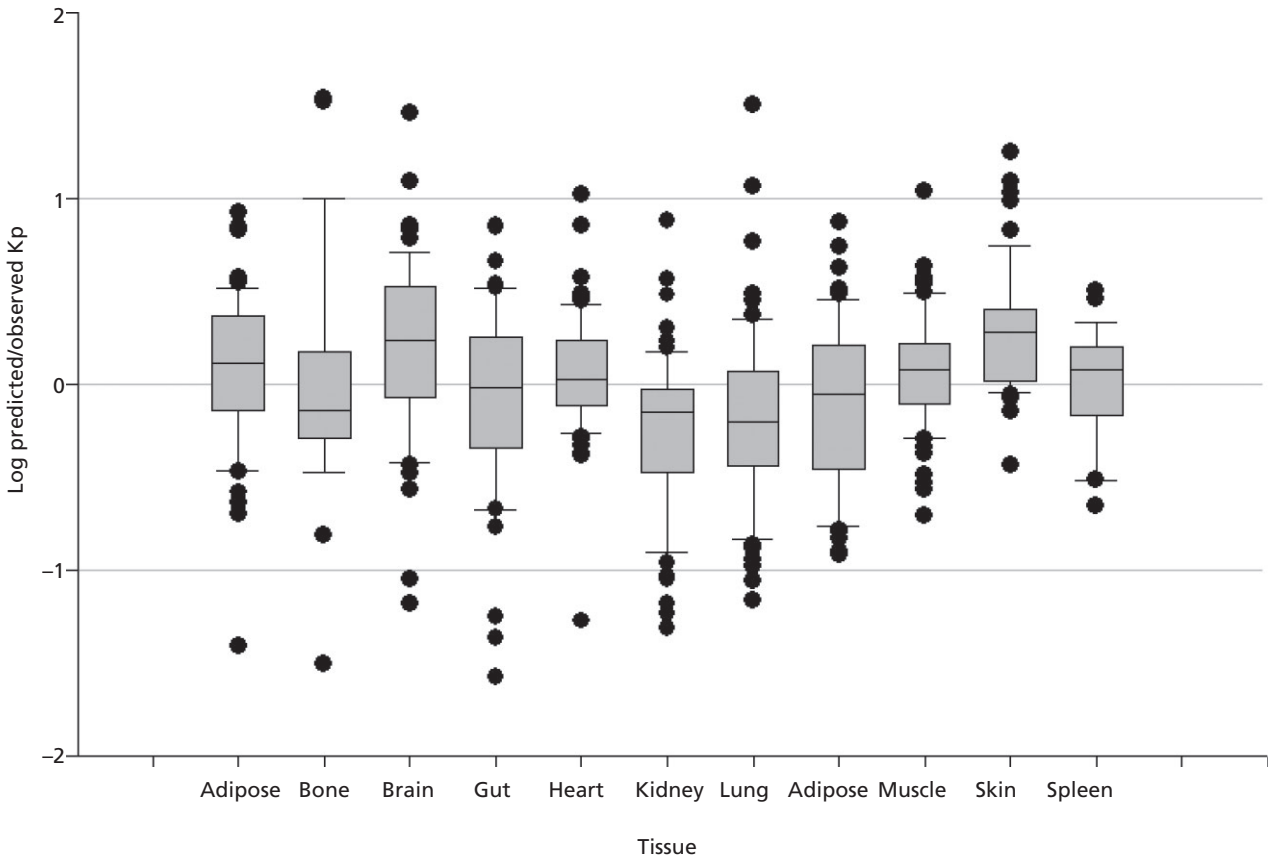

Figure 3 Boxplot of log predicted/observed *K*p values for each tissue as predicted by the Rodgers *et al*. model.[9,17] The boxes indicate the standard deviation; the whiskers represent the range between the 10th and 90th percentile; the horizontal black line represents the median value; and outliers are represented by black circles.

66.1% in brain to 92.7% in heart). The Poulin & Theil model was the least consistent, with a range of 77.1% (20.0% in spleen to 97.1% in heart).

Outliers in Kp *prediction*

Large under predictions (by up to a factor of 100) were seen for acebutolol-R and acebutolol-S in gut tissue in all models, and betaxolol-R and betaxolol-S were under-predicted in lung tissue by all models except the Poulin & Theil model. Propranolol-R and propranolol-S were poorly predicted in brain tissue (under-predicted by up to a factor of 16) by all except the Arundel and Poulin & Theil models. The single largest prediction error was shown by the Berezhkovskiy model for betaxolol-R in the lung, which was under-predicted by a factor of 224.

V_{ss} predictions

Of the four models for which V_{ss} predictions were made, the Poulin & Theil model was the most accurate, with 87.0% of predictions within threefold of experimental values (Figure 4; Table 4). However, there was no significant difference in the accuracy compared with the Rodgers *et al*. model that resulted in slightly less number of studies within that range (80.3% within threefold, $P = 0.1$). The Berezhkovskiy and Poulin *et al*. models were the least accurate, both with 63.0% of predictions within the threefold range, and this result was significant ($P < 0.05$). When experimental *K*p values were used to predict V_{ss}, 88.9% of predictions were within threefold of experimental values, which was significantly more accurate than Rodgers *et al.*, Berezhkovskiy, and Poulin *et al.* models ($P < 0.5$); no statistically significant difference was observed when compared with the performance of the Poulin & Theil model ($P = 0.5$).

Discussion

*K*p predictions

The current study assessed both empirical and mechanistic models for their ability to predict partition coefficients of 81 drugs across a range of tissues. The comparisons performed in this study showed that the Rodgers *et al.* model is the

significantly the most accurate as a general partition coefficient predictor across all drug classes and all tissues, with more than 77% of the predictions within threefold of experimentally derived values. This model was consistently shown to be the most accurate across all tissue types. The Berezhkovskiy and Poulin *et al*. models were significantly the least accurate, with > 30% of predictions outside fivefold of experimental values.

Assessing the results for the individual drug classes indicated that the Poulin *et al*. model performed particularly poorly for basic compounds, with less than 42% of predictions within the threefold range. Poulin *et al*. themselves acknowledged this deficiency of their model, and it was for this reason that the Rodgers *et al*. model concentrated specifically on improving the accuracy of predictions for moderate-to-strong bases in response to the Poulin *et al*. papers. Rodgers *et al*. incorporated the unique binding properties of these types of bases into their model, and therefore it is unsurprising that this model performs much better in the prediction of *K*p values for compounds of this type. Although the Poulin *et al*. and Rodgers *et al*. models are both mechanistic in nature, the Rodgers *et al*. model incorporates a greater number of potential interactions with tissue components which may explain the improvement in accuracy of prediction exhibited by this model. The Rodgers *et al*. model takes into account not only the volumes of neutral lipids and phospholipids in the tissue, but also the interactions with acidic phospholipids and also the potential specific binding of a compound to extracellular proteins such as albumin and lipoproteins. For moderate-to-strong bases, the incorporation of drug dissolution in the tissue water and the partitioning of unbound, un-ionised drug into neutral lipids and neutral phospholipids also lead to an improvement in accuracy when compared with earlier models that neglect to include these important and influential features.

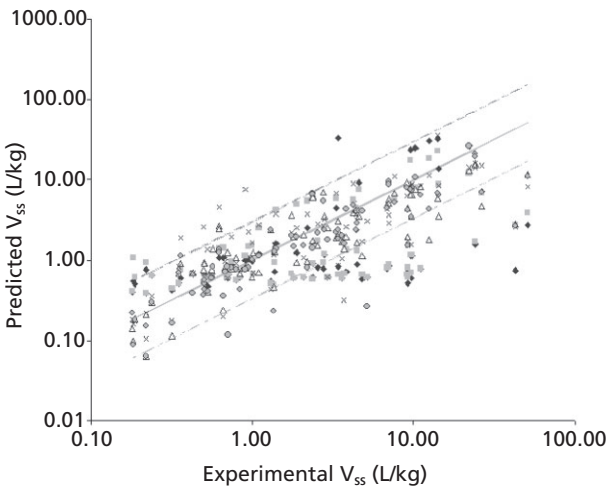

Figure 4 Relationship between predicted and experimentally derived V_{ss} (L/kg) values for 81 compounds using the Berezhkovskiy model (diamonds), the Poulin *et al*. model (squares), the Rodgers *et al*. model (crosses), the Poulin & Theil model (triangles), and experimental *K*ps (circles). Straight line indicates line of unity and dashed lines indicate threefold above and below.

The Poulin *et al*. and Berezhkovskiy models were designed primarily to predict *K*p for acids and weak bases, and therefore they show a low degree of accuracy for the basic compounds in this dataset, as 37 of the 46 bases in the dataset are classed as strong (i.e. pKa $\geq$ 7). As expected, these two models show much better accuracy of prediction for the acidic compounds in this dataset, with over 71% of predictions within the threefold range. The dataset itself is biased towards basic compounds, containing 46 bases compared with only 21 acids, and so this may have skewed the results in favour of certain models. However, by looking at the results for the different drug classes separately alongside the overall results, it is possible to remove the influence of this bias and further understand where the strength of each model lies.

A threefold range was chosen to measure the accuracy of the predictions made by each model as it has been shown previously that experimental values can differ due to inter-laboratory and inter-animal variability.[15] Therefore, to provide a definitive validation of the prediction ability of a model, experimental data for all compounds would need to be gathered from the same source. Using only data available in the literature, this is an impossible task for such a large dataset. The experimental values collated from the literature for this study come from a large variety of different sources

Table 4 Comparative and statistical assessment of the four models used to predict V_{ss} in rat using 81 compounds

Model	% <3-fold	% 3–5-fold	% >5-fold	afe	aafe	rmse	ccc	*n*	mean pred : obs ratio	SD
Berezhkovskiy	63.0	11.1	25.9	0.55	2.87	0.60	0.26	81	1.03	1.27
Poulin *et al*.	63.0	11.1	25.9	0.58	2.85	0.57	0.26	81	1.04	1.04
Rodgers *et al*.	80.3	5.3	14.5	0.92	2.04	0.42	0.45	76	1.43	1.60
Poulin & Theil	87.0	9.1	3.9	0.78	1.84	0.35	0.45	77	1.02	0.80
from experimental *K*ps	88.9	4.9	6.2	0.74	1.69	0.34	0.53	81	0.91	0.59

aafe, absolute average fold-error; afe, average fold-error; ccc, correlation concordance coefficient; obs, observed; pred, predicted; rmse, root mean squared error; SD, standard deviation.

and therefore are likely to display a degree of variability that is impossible to quantify. Use of the threefold factor as a measure of accuracy allowed the effect of this variability to be accounted for to some extent. Wherever possible, the in-vivo *K*p data used in this study was obtained under steady-state or pseudo-equilibrium conditions; however, for some studies no data were available to confirm this.

Another degree of uncertainty is introduced by considering the variability of the input parameters used by the models in this study. The sensitivity of a particular model to one of these input parameters can be investigated by varying that parameter and examining the effect upon the results. For example, the logP values used here were experimentally determined wherever possible (65 out of the 81 compounds); however, when only predicted logP values were used[9,12,17] (details not shown here), the number of *K*p predictions within threefold made by the Rodgers *et al*. model was reduced to 71%, a significant decrease in accuracy of almost 7% ($P < 0.01$). Furthermore, increasing all logP values by just 20% caused a significant decrease in accuracy of the Rodgers *et al*. model of almost 14% ($P < 0.01$). Repeating the same procedure with pKa values showed a significant reduction in accuracy of the Rodgers *et al*. model of almost 9% ($P < 0.05$) when the input parameter is varied by $\pm$ 20%. This highlights the need for caution when interpreting any results that rely on such a large amount of experimental data from different sources.

Certain drugs in certain tissues were predicted poorly by nearly all models, such as acebutolol in gut, bone and brain, betaxolol in lung, and propranolol in brain. The reasons for these poor predictions are likely to be complex and may involve a variety of mechanisms, and some of these issues have been highlighted previously and commented upon in the literature. For example, Rodgers *et al*. also observed the over-prediction of acebutolol in gut and discussed the possibility that this may arise from analysing the gut along with its contents when deriving the in-vivo *K*p value. Processes such as intestinal reabsorption and biliary secretion could be contributing to an increase in concentration of compound in the gut contents and therefore a true comparison cannot be made between predicted results (gut tissue only) and experimental results (gut tissue plus gut contents).[20] The authors also saw a degree of under-prediction for certain compounds in lung, which is also seen in this study. This could be attributed to the possibility that basic, lipophilic compounds might be sequestered within lysosomes in the lung tissue.[21,22] It is also known that other tissues, such as liver and kidney, are rich in lysosomes and yet the same 'lysosomal trapping' effect as seen in lung is not observed. This goes to further show that the partitioning of drug into these organelles is little understood, and therefore it is not yet possible to incorporate this mechanism into predictive models.

However, it is worth noting that all of the experimental *K*ps for the three outliers, acebutolol, propranolol and betaxolol, were taken from the same study.[20] Although the authors took precautions to ensure that all *K*p values were obtained under steady-state conditions (including comparison of results from 4-h and 8-h infusions and analysing all results in triplicate), it is possible that some errors were made.

Another major limitation of all the models investigated in this study is their lack of incorporation of active uptake processes, which can greatly affect the tissue distribution of certain compounds.[23] All the models assume that compounds are distributed solely by passive diffusion leading to perfusion-limited distribution and minimal contribution from active uptake. Pravastatin is an example of a drug that undergoes active uptake into both the liver and the kidney via the organic anion transporters OATP1B1 and OATP3, respectively.[24,25] If predicted *K*p values for this compound are compared to in-vivo *K*p data reported for a minimal number of tissues,[26] the Rodgers *et al*. model significantly under-predicts *K*p for kidney as the predicted value represents only 3.5% of the observed *K*p (19.4). Use of empirical models results in even more pronounced inaccuracy, as the predicted *K*p for this tissue represented only 0.5–2% of the observed value using the Arundel and Jansson *et al*. models, respectively. If, in contrast, *K*p is predicted for a tissue not known to actively take up pravastatin (e.g. lung), the models produce more accurate results, with the Rodgers *et al*. model predicting *K*p within twofold of experimental values, and the Jansson *et al*. model under-predicting *K*p by less than threefold. These findings indicate the limitations of the current models to accommodate active uptake processes and this, consequently, has implications on the application of these predictive tools for PBPK modelling of such compounds in the drug development process. Considering the increasing number of drugs found to be associated with active uptake, in particular in the liver (e.g. statins, repaglinide),[27,28] refinement of current predictive models and incorporation of the permeability limited distribution is required.

V_{ss} predictions

If a specific *K*p value for a certain tissue was required in the drug development process, it is likely that it would be determined experimentally rather than relying on a prediction method such as the ones shown here. Therefore, the most widespread use of *K*p prediction is not in identifying individual values but in using the *K*p values of all the tissues together to generate a prediction for the volume of distribution at steady state (V_{ss}) of a compound. Therefore, this study also investigated the accuracy of V_{ss} predictions made using the predicted *K*p values from four of the models. The Poulin & Theil model was shown to be the most accurate, with 87% of predictions within threefold of experimental values. As muscle tissue makes up more than 50% of the total tissue volume in rat, accurate *K*p predictions for this tissue would be

expected to lead to accurate V_{ss} predictions. As the Poulin & Theil model does not predict muscle *Kp*, and instead uses in-vivo muscle *Kp* as an input parameter, it is to be expected that this model provides the most accurate V_{ss} predictions. Of the other three models that do predict muscle *Kp*, the Rodgers *et al*. model performs better than the other models, with over 83% of muscle *Kp* predictions within threefold of experimental values. This explains to some extent why the Rodgers *et al*. model also provides the most accurate V_{ss} predictions out of these three models, with over 80% within the threefold range.

Previous studies looking at the prediction of human pharmacokinetics have shown that 78%,[13] 79%[14] and 79%[16] of V_{ss} predictions made by the Rodgers *et al*. model are within threefold of experimental values, although much smaller datasets were used in these analyses (18, 26 and 47 compounds, respectively). The authors of the Poulin & Theil model showed 89.4% of predictions made by their model were within threefold of experimental values for a dataset that consisted solely of basic compounds. Consequently, these results from the literature are comparable with the results from this study, despite this study focusing solely on rat V_{ss} and the other studies on human values. However, for the other two models investigated here this was not the case. The recent Jones *et al*.[13] study found that for the Poulin *et al*. model, 61% of predictions were within threefold of experimental values, compared with 51% in this study. For the Berezhkovskiy model, the difference is more pronounced, with values of 72% and 46%, respectively. This can perhaps be explained by the differences in the datasets used by the two studies. In the Jones *et al*. study, 55% ($n = 6$) of the basic compounds can be classified as strong bases (pKa $\geq$ 7), whereas in the dataset used here, almost 80% ($n = 37$) of the basic compounds fit this criterion. As described above, the Poulin *et al*. and Berezhkovskiy models have previously been shown to perform poorly in the prediction of *Kp* for this type of compound, and so are likely to produce better V_{ss} predictions for the Jones *et al*. dataset, which contains a higher proportion of weak basic compounds. The dataset used here also covers a much wider range of physicochemical properties than the Jones *et al*. study, with a 1.7-fold and 1.9-fold larger range of pKa and logP values, respectively. Therefore, it is unlikely that exactly comparable results would be produced by the two studies.

Unexpectedly, the number of accurate V_{ss} predictions made using experimental *Kps* was not significantly larger than the Poulin & Theil model, with just under 90% of predictions within threefold of experimental values. This result is due, in part, to a lack of experimental data for certain compounds. For example, for morphine, experimental *Kp* values were only available in gut, kidney and liver, and so the predicted V_{ss} value was 19-fold smaller than the experimental value. As these three tissues only make up 10% of the rat tissue volume, it is unsurprising that V_{ss} predictions made from them are highly inaccurate. Conversely, the V_{ss} prediction for theophylline, for which again only three experimental *Kps* were available, is within twofold of experimental V_{ss}. *Kps* for this drug are available in brain, muscle and lung, which make up 58% of the rat tissue volume, and so more accurate V_{ss} predictions are possible. Caffeine has five experimental *Kp* values available (adipose, bone, brain, heart, kidney; 14% of tissue volume), and the V_{ss} prediction is more than six times smaller than the experimental value.

Conclusions

In conclusion, the Rodgers *et al*. model has been shown to be the most accurate a-priori model for the prediction of both *Kp* and V_{ss} values in rat. However, the model does show enough limitations to justify the further development and improvement of this method to increase its reliability and allow it to be used with more confidence during the drug development process. The incorporation of elements such as active transport systems and lysosomal trapping is required to further enhance the accuracy of these models, although this could potentially lead to the models becoming too complex and too specific to be of general use.

Declarations

Conflict of interest

The Author(s) declare(s) that they have no conflicts of interest to disclose.

Funding

One of the authors (Helen Graham) thanks the BBSRC and AstraZeneca, Alderley Edge, UK for their financial support in the form of a CASE studentship.

References

1. Edginton AN *et al*. Whole body physiologically-based pharmacokinetic models: their use in clinical drug development. *Expert Opin Drug Metab Toxicol* 2008; 4: 1143–1152.
2. Charnick SB *et al*. Perspectives in pharmacokinetics. Physiologically based pharmacokinetic modeling as a tool for drug development. *J Pharma-*

cokinet Biopharm 1995; 23: 217–229.
3. von Kleist M, Huisinga W. Physiologically based pharmacokinetic modelling: a sub-compartmentalized model of tissue distribution. *J Pharmacokinet Pharmacodyn* 2007; 34: 789–806.
4. Björkman S. Prediction of the volume of distribution of a drug: which tissue-plasma partition coefficients are needed? *J Pharm Pharmacol* 2002; 54: 1237–1245.
5. Poulin P, Krishnan K. An algorithm for predicting tissue: blood partition coefficients of organic chemicals from n-octanol: water partition coefficient data. *J Toxicol Environ Health* 1995; 46: 117–129.
6. Nestorov IA *et al*. Quantitative structure-pharmacokinetics relationships: II. A mechanistically based model to evaluate the relationship between tissue distribution parameters and compound lipophilicity. *J Pharmacokinet Biopharm* 1998; 28: 521–545.
7. Parham FM *et al*. Using structural information to create physiologically based pharmacokinetic models for all polychlorinated biphenyls: I. Tissue:blood partition coefficients. *Toxicol Appl Pharmacol* 1997; 144: 340–347.
8. Poulin P *et al*. Prediction of adipose tissue: plasma partition coefficients for structurally unrelated drugs. *J Pharm Sci* 2001; 90: 436–447.
9. Rodgers T *et al*. Physiologically based pharmacokinetic modeling 1: predicting the tissue distribution of moderate-to-strong bases. *J Pharm Sci* 2005; 94: 1259–1276.
10. Berezhkovskiy LM. Volume of distribution at steady state for a linear pharmacokinetic system with peripheral elimination. *J Pharm Sci* 2004; 93: 1628–1640.
11. Arundel PA. A multi-compartmental model generally applicable to physiologically-based pharmacokinetics. 3rd IFAC Symposium, University of Warwick, UK, March 23–26, 1997.
12. Jansson R *et al*. Prediction of drug tissue to plasma concentration ratios using a measured volume of distribution in combination with lipophilicity. *J Pharm Sci* 2008; 97: 2324–2339.
13. Jones RD *et al*. PhRMA CPCDC initiative on predictive models of human pharmacokinetics, part 2: comparative assessment of prediction methods of human volume of distribution. *J Pharm Sci* 2011; 100: 4074–4089.
14. De Buck SS *et al*. Prediction of human pharmacokinetics using physiologically based modeling: a retrospective analysis of 26 clinically tested drugs. *Drug Metab Dispos* 2007; 35: 1766–1780.
15. Poulin P, Theil FP. A priori prediction of tissue:plasma partition coefficients of drugs to facilitate the use of physiologically-based pharmacokinetic models in drug discovery. *J Pharm Sci* 2000; 89: 16–35.
16. Poulin P, Theil FP. Development of a novel method for predicting human volume of distribution at steady-state of basic drugs and comparative assessment with existing methods. *J Pharm Sci* 2009; 98: 4941–4961.
17. Rodgers T, Rowland M. Physiologically based pharmacokinetic modelling 2: predicting the tissue distribution of acids, very weak bases, neutrals and zwitterions. *J Pharm Sci* 2006; 95: 1238–1257.
18. Nestorov IA *et al*. Lumping of whole-body physiologically based pharmacokinetic models. *J Pharmacokinet Biopharm* 1998; 26: 21–46.
19. Leo A *et al*. Partition coefficients and their uses. *Chem Rev* 1971; 71: 525–616.
20. Rodgers T *et al*. Tissue distribution of basic drugs: accounting for enantiomeric, compound and regional differences amongst beta-blocking drugs in rat. *J Pharm Sci* 2005; 94: 1237–1248.
21. MacIntyre AC, Cutler DJ. The potential role of lysosomes in tissue distribution of weak bases. *Biopharm Drug Dispos* 1988; 9: 513–526.
22. Hallifax D, Houston JB. Saturable uptake of lipophilic amine drugs into isolated hepatocytes: mechanisms and consequences for quantitative clearance prediction. *Drug Metab Dispos* 2007; 35: 1325–1332.
23. Berezhkovskiy LM. The influence of hepatic transport on the distribution volumes and mean residence time of drug in the body and the accuracy of estimating these parameters by the traditional pharmacokinetic calculations. *J Pharm Sci* 2011; 100: 5031–5047.
24. Yamazaki M *et al*. Tissue-selective uptake of pravastatin in rats: contribution of a specific carrier-mediated uptake system. *Biopharm Drug Dispos* 1996; 17: 775–789.
25. Kusuhara H, Sugiyama Y. In vitro-in vivo extrapolation of transporter-mediated clearance in the liver and kidney. *Drug Metab Pharmacokinet* 2009; 24: 37–52.
26. Watanabe T *et al*. Physiologically based pharmacokinetic modeling to predict transporter-mediated clearance and distribution of pravastatin in humans. *J Pharmacol Exp Ther* 2009; 328: 652–662.
27. Niemi M *et al*. Polymorphic organic anion transporting polypeptide 1B1 is a major determinant of repaglinide pharmacokinetics. *Clin Pharmacol Ther* 2005; 77: 468–478.
28. Ieiri I *et al*. Genetic polymorphisms of uptake (OATP1B1, 1B3) and efflux (MRP2, BCRP) transporters: implications for inter-individual differences in the pharmacokinetics and pharmacodynamics of statins and other clinically relevant drugs. *Expert Opin Drug Metab Toxicol* 2009; 5: 703–729.
29. Schmitt W. General approach for the calculation of tissue to plasma partition coefficients. *Toxicol In Vitro* 2008; 22: 457–467.
30. Mostafavi SA, Foster RT. Pharmacokinetics of acebutolol enantiomers after intravenous administration of racemate in a rat model: a dosing range comparison. *Biopharm Drug Dispos* 1997; 18: 397–408.
31. Tahara K *et al*. Pharmacokinetics and pharmacodynamics of bisoprolol in rats with bilateral ureter ligation-induced renal failure. *Drug Metab Pharmacokinet* 2006; 21: 389–394.
32. Rodgers T, Rowland M. Mechanistic approaches to volume of distribution predictions: understanding the processes. *Pharm Res* 2007; 24: 918–933.
33. Klotz U. Estimation of the blood-plasma concentration ratio of

diazepam in the rat. *J Pharmacokinet Biopharm* 1985; 13: 347–348.
34. Cohen AM. The disposition of ftorafur in rats after intravenous administration. *Drug Metab Dispos* 1975; 3: 303–308.
35. Kawai R *et al*. Physiologically based pharmacokinetic study on a cyclosporin derivative, SDZ IMM 125. *J Pharmacokinet Biopharm* 1994; 22: 327–365.
36. Poulin P, Theil FP. Prediction of pharmacokinetics prior to in vivo studies. 1. Mechanism-based prediction of volume of distribution. *J Pharm Sci* 2002; 91: 129–156.

Supporting Information

Additional Supporting Information may be found in the online version of this article:

Table S1 In-vitro data for use as input parameters in models used to predict *K*p values in the rat
Table S2 Rat tissue composition data for use as input parameters in models used to predict *K*p values
Table S3 Rat *K*p predicted via six published equations followed by experimentally determined *K*p
Table S4 Rat V_{ss} (L/kg) predicted via four *K*p prediction methods and by experimental *K*ps, followed by experimentally determined V_{ss} (L/kg)

Please note: Wiley-Blackwell are not responsible for the content or functionality of any supporting materials supplied by the authors. Any queries (other than missing material) should be directed to the corresponding author for the article.